‘What does the cow say?’

An exploratory analysis of onomatopoeia in early phonological development

Catherine Elizabeth Laing

PhD

Language and Linguistic Science
University of York

October 2015

ABSTRACT

This thesis presents an in-depth analysis of infants' acquisition of onomatopoeia – an area of phonological development that until now has been largely overlooked. Infants produce many onomatopoeia in their earliest words, which are often disregarded in phonological analyses owing to their marginal status in adult languages. It is often suggested that onomatopoeia may be easier for infants to learn because of the iconicity that is present in these forms; this corresponds to Imai and Kita's (2014) 'sound symbolism bootstrapping hypothesis', as well as Werner and Kaplan's theoretical work *Symbol Formation* (1963). However, neither of these accounts considers the role of phonological development in infants' acquisition of onomatopoeia.

This thesis presents a series of six studies with a range of perspectives on our central research question: is there a role for onomatopoeia in phonological development? Two analyses of longitudinal diary data address the nature of onomatopoeia in early production, while two eye-tracking studies consider the nature of iconicity in onomatopoeia and whether or not this has a perceptual advantage in early development. The role of the caregiver is then considered, with a prosodic analysis of onomatopoeia in infant-directed speech and a longitudinal perspective of the role of onomatopoeia in infant-caregiver interactions.

The contributions from thesis are threefold. First, we offer empirical evidence towards an understanding of how onomatopoeia fit within an infant's wider phonological development, by showing how onomatopoeia facilitate early perception, production and interactions. Second, our results illustrate how these forms are an important aspect of phonological development and should not be overlooked in infant language research, as has often been the case in the development literature. Finally, these findings expand the iconicity research by showing that onomatopoeia do not present an iconic advantage in language learning, as has so often been assumed by theorists in the field.

Table of contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I consider myself extremely lucky to have had such a fulfilling PhD experience, and this can only be attributed to those who I will mention below. These pages constitute the physical result of three years of hard work, but I must also acknowledge those aspects of the PhD that can't be held in the hands – I'm sure I will carry them with me long after this thesis is awarded a space on my shelf.

First, I am indebted to the Economic and Social Research Council, which has funded my research over the past three years. Alongside the grant that allowed me to live and study without compromise, I was also lucky enough to take on an Overseas Institutional Visit to Stockholm University thanks to their generosity, as well as attending conferences and running the PENCIL project. All I can say is thank you – this has been the best experience of my life so far.

Similarly, I want to thank the NorPhLex group for putting on an exciting programme of workshops and conferences (in equally exciting places!) and for funding my attendance at each of these. Much of the research in this thesis and many of the best experiences of my PhD have come as a result of my first NorPhLex experience in Stockholm, and many of the members of the group have now become good friends. This leads me to Stockholm University itself, where I spent a very happy month learning about eye-tracking. Thank you to Francisco Lacerda for allowing it to happen, to Iris-Corinna Schwartz for inspiring me to make it happen, and to Lena Renner, Ellen Marklund, Elizabet Eir Cortes, Petter Kallioinen, Marie Markelius and Ulla Sundberg for all of the assistance, support and friendship that you provided during that period.

I am also very lucky to have had the support of some great colleagues in York, without whom it would not have been possible to run my eye-tracking experiment. Thank you to Huw Llewelyn-Jones for the technical support, and for averting a couple of crises which still make my stomach turn today. Thanks to Mariam Dar, too: aside from helping recruit participants and showing me the Babylab 'ropes', you have been a brilliant source of moral support over the past two years in particular, and I am very glad of our friendship. Ben Miller, next-door-neighbour extraordinaire, deserves a special mention – thanks for saving the day when my experiment wouldn't run; I owe you much more than the cake that you received for all of your help.

York's Department of Language and Linguistic Science is a fantastic place to be a PhD student, and many staff members have helped me in some way over the past three years. In particular I want to mention Sam Hellmuth and Richard Ogden, who have both helped to shape some of the thoughts that have found their way into this thesis. Thanks also must go to Virve Vihman, Rory DePaolis, Tamar Keren-Portnoy and Marilyn Vihman for allowing me to use their data in my analyses, and also to Márton Sóskuthy and Christoph Scheepers for taking the time to go over some of my statistics.

Of course, this adventure would never have begun without Marilyn Vihman and Tamar Keren-Portnoy. Both of you taught the undergraduate Phonological Development module that unexpectedly became my MA, and you have continued to inspire, motivate and challenge me throughout my PhD. Every supervision meeting has left me feeling upbeat and motivated, and your confidence in me has helped to fire up confidence in myself, even in some of the more difficult moments. Thank you so much for everything: both your tutelage and your company has been second to none.

When I started this project I never considered that I might make a few friends along the way, and I certainly didn't expect to find two of the best friends I've ever had. Jessica Wormald and Claire Canavan – we've gone through a lot of this experience together, and sharing the ups and downs with you both has been an absolute pleasure. Thanks for the many occasions that you dragged me out of bed at 6am for some fresh air and a natter; thanks for running alongside me – I hope it will continue for many years to come.

And now it comes to my nearest and dearest. Mum and Dad, who have supported me through everything, and who have cheered me on so proudly over the past three years. Also my brother Will, who always helps me keep things in perspective. I can't list all of the names, but I want to dedicate this thesis to all of my family, past and present. Finally, to Daniel, who first inspired me to do a PhD (I didn't know what a PhD was until we met), and who has been the ultimate source of support, love and comfort. You have celebrated and commiserated alongside me throughout, and there is nothing more reassuring that knowing that I get to do all of my celebrating and commiserating with you. You and Edward are the best support team I could ask for – thank you.

DECLARATION

I declare that the content of this thesis is my own work, and any content from external sources has been acknowledged through explicit referencing and quotations where appropriate. No part of this thesis has been submitted for examination at this or any other institute for another award.

I also declare that parts of this thesis have been previously published in journals and conference proceedings, or are currently under review. These publications are as follows:

- Laing, C. E. (under review). Prosodic trends in infants' acquisition of onomatopoeia: An analysis of the first 100 words.

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

- Laing, C. E. (2014). Phonological 'Wildness' in Early Language Development: Exploring the Role of Onomatopoeia, *Proceedings from the PARLAY Conference 2013, York Papers in Linguistics*.

Findings from Chapter 6 are currently under review. I was the lead author on this paper, and independently responsible for all analyses. However, data was collected by the second and third authors, and is cited as such in Chapter 6 of this thesis.

- Laing, C. E., Keren-Portnoy, T. & Vihman, M. M. (under review). The role of onomatopoeia in language development: A prosodic analysis of infant-directed speech.

If we lived in a world where bells
truly say 'ding dong' and where 'moo'
is a rather neat thing said by a cow,
I could believe you could believe
that these sounds that we make in the air
and these shapes with which I blacken white paper
have some reference to the thoughts in my mind
and the feelings in the thoughts.

- from *Linguist* by Norman MacCaig

1. INTRODUCTION

1.1 Overview

Since the first considerations of the emergence of human speech, it has been proposed that onomatopoeia may be central to the evolution of language. Onomatopoeia constitute a subset of words that are classed as 'iconic', defined by the word's constituents – from the full lexical unit to the individual phonemes – which somehow represent its lexical meaning. Going as far back as the Ancient Greeks, the question of a word's form-meaning relationship has been considered largely from two main perspectives: Plato's *Cratylus* comprises an extensive discussion of the nature of language (Barney, 2001), whereby Cratylus and Hermogenes debate the semantic importance of the symbol-referent relationship. Cratylus defends the notion of iconicity and the "natural fitness of names" (Jowett, 1953, 391a), while Hermogenes argues that a name is merely "a portion of the human voice which men agree to use" (383a). Here we find the earliest and possibly the most extensive discussion of what has come to be known as 'the arbitrariness of the sign' (de Saussure, 1962; original publication 1915) – a debate which continues over two thousand years later.

The discussion in *Cratylus* eventually moves towards onomatopoeia, when Socrates questions the expression of names through imitation:

> "[the body] must imitate the thing to which it would refer…so when we want to express something with the voice…the expression will be achieved by imitation…of that which we want to express?...[Then] we shall be obliged to admit that the people who imitate sheep, or cocks, or other animals, *name that which they imitate*" (Jowett, 1953, 423a-c, italics added).

Socrates soon dismisses this idea as "wild and ridiculous" (426b), but in fact this early debate regarding the establishment of form-meaning correspondences – with the *semantic method* (Hermogenes' argument) at one end of the spectrum and the *mimetic method* (Cratylus' argument; Baxter, 1992) at the other – remains active today. Indeed, while de Saussure's (1915/62) analysis of arbitrariness existed as the accepted view on this matter for almost a century (but see Köhler, 1970, for a contrasting perspective), a recent resurgence in empirical studies in this area has allowed us to return to the possibility that words and their referents may not be quite as arbitrary as Hermogenes suggests (Asano et al., 2015; Monaghan, et al., 2014; Ozturk at el., 2013).

Even in the past sixty years it has been proposed that infants 'create' words through the imitation of natural sounds, establishing an early lexicon of onomatopoeic protowords that enables infants to refer to the objects around them (Werner and Kaplan, 1963). What was once considered a "wild and ridiculous" (Jowett, 1953, 426b) or even a merely outdated view of language phylogeny has returned as an accepted view in psycholinguistic research, with numerous experimental studies vouching for the advantage of iconic forms in infant word learning (Asano et al., 2015; Kantartzis et al., 2011). So far this discussion has remained largely one-sided, with a wealth of studies focussing on the perception of 'sound symbolic' forms, whereby certain phonemes are thought to represent physical properties, such as /i/ to represent sharp or spiky objects and /o/ to reflect round or bulbous shapes. Advanced experimental methods have been used to test both adults and infants' responses to sound symbolic words, with the general conclusion that iconicity presents an advantage to language learners of all ages. However, few studies have acknowledged the early lexicon and the role of production – an analysis of which could open the door to further understanding in this field. Indeed, onomatopoeia are a common feature of infants' earliest words (Menn & Vihman, 2011, Appendix), and a consideration of these forms in terms of both perception and production could provide a new and balanced perspective on this centuries-old argument.

In this thesis we will explore the nature of these less-than-arbitrary, or 'iconic', forms – defined by Wescott (1971) as words which are "imitative of non-linguistic reality" (p.416) – and their potential role in the emergence of early infant language. The acquisition of onomatopoeia – here defined as the linguistic approximation of a natural sound in reference to an object or concept – has remained unexplored in this field, despite ample discussion and research in both child language acquisition and sound symbolism occurring in parallel over the course of the past century. Since the Saussurean notion of arbitrariness was posited in the early 1900s, empirical research has shown us that, across languages, humans do indeed make consistent sound-symbol correspondences (Köhler 1970) which point to the existence of deeply embedded mimetic understanding in our perception of language. It has been suggested that this may be an innate aspect of primate cognition (Ozturk et al., 2013), and the question of whether sound-form correspondences might be advantageous in infant language learning has recently become a popular topic for investigation in the language-learning literature (Asano et al., 2015; Imai & Kita, 2014; Kantartzis et al., 2011; Monaghan et al., 2014); this goes against the traditional view that, for the most part, language is arbitrary (de Saussure, 1915/62). However, onomatopoeia are yet to be considered from this perspective, despite being acknowledged as one of the few examples of 'true' sound symbolism (Lyons, 1968; Sapir, 1970), and despite their presence in a wide range of the world's languages (Hinton, Nichols & Ohala, 1994), if not universally.

1.2 Sound symbolism: the current perspective

Although a dichotomous view similar to that of the Cratylus versus Hermogenes debate (Jowett, 1953) has largely been presented so far, in fact it is now widely accepted that, in terms of the presence of iconicity, most human languages fall somewhere between these two extremes, with some displaying a higher frequency of symbolic words than others. Japanese and Korean, as well as many other Southeast Asian and sub-Saharan languages (Imai & Kita, 2014), possess a rich sub-lexicon of symbolic words, or mimetics, which function as an integral part of formal speech, and which are reported to be notoriously difficult for non-native speakers to learn (Ivanova, 2006). Consequently, Japanese features heavily in many discussions and studies in this area, with its lexicon of around 4,500 mimetic words (Imai & Kita, 2014) providing ample material for analysis. At the other end of the spectrum, most Indo-European languages have relatively few examples of iconicity, though still a certain number of less-than-arbitrary forms can be found, with the famous example of phonaesthetic 'gl-words' in English, such as *glisten*, *glitter*, *gleam* and *glimmer*, which all relate to a similar concept of twinkling light (Bergen, 2004). It has been proposed that these phonoaesthemes function as meaningful units in language, with semantic properties similar to those of morphemes (Magnus, 2001). However, Bergen (2004) posits that the 'psychological status' of these phonological segments can be attributed to the statistical regularities of sound-meaning pairings in the input.

Typically, discussions on sound symbolism relate to what Ohala (1984) termed the 'frequency code'; that is, a symbolic relationship between the formant values and vocal tract size in the production of a specific segment, and its corresponding meaning. Vowels and consonants with a smaller vocal tract size, and thus higher f0, such as /i/ and /k/, refer to small, sharp or rapid referents, while those produced with a larger space in the vocal tract and lower f0 such as /u/ and /b/ relate to large, slow or heavy referents (Hinton et al., 1994). Mimetics are also derived from these correspondences (Ivanova, 2006); Kita (1997, p.380) claims that "in [the] realm of mimetic forms, phonemes seem to have meanings of their own". This segmental relationship between sounds and meanings is drawn upon in a wide range of studies, most famously in Köhler's (1970) experiments involving the 'bouba-kiki effect'. Here, Köhler showed how participants consistently established meaningful correspondences between specific phonological segments and rounded or angular shapes. Evidence shows these correspondences to be consistent across different languages and age-ranges (Davis, 1961), as well as when participants are presented with both novel and familiar objects (D'Onofrio, 2014) or words in foreign languages (Imai et al., 2008; Nygaard, Cook & Namy, 2009). However, there is contrasting evidence regarding the phonological manifestation of sound symbolic features: Maurer, Pathman and Mondloch (2006), for example, show that vowel rounding determines the symbolic features of a form, while Fort and colleagues (2014) found a stronger role for consonants here. Furthermore, Ozturk and colleagues (2013) found that both vowels and consonants provided sufficient sound symbolic information in their own right to enable adults to map particular sounds to particular object properties, though four-month-old infants required congruity across both vowels and consonants (i.e. both 'high' as in *kiki* or both 'low' as in *bubu*) to succeed in this task. D'Onofrio (2014) explored the phonological specification of this effect in even further depth, finding that voicing, place of articulation and vowel backness had an effect on participants' judgements of which segmental features constituted a round shape or an angular shape. Similar results were observed when the same analysis was carried out on real-world objects ('round' objects such as a bowl or a spoon and 'angular' objects such as a knife or a cheese-grater).

In a fascinating analysis of perception in synaesthesia, Ramachandran and Hubbard (2001) interpret the neuropsychological experience of sound symbolism. Their account discusses the evolution of language as a coming-together of various sensory systems, which prioritise sound symbolic correspondences owing to their multi-sensory networks (pairings between particular sound contours and corresponding object shapes, articulatory movements and gestures) being stronger than single-modality connections. This co-activation of multiple networks is briefly alluded to with reference to mirror neurons, and relates to the combined somatosensory experiences that occur when the articulation of a segment takes place alongside the perception of an object of a specific shape or size. For example, the production of a rounded vowel, which necessarily brings about the rounding of the lips, may enable speakers to establish multimodal connections with an object which is also visually rounded, as oppose to the single-modality connection between the same rounded vowel and a shape which is angular. Evidence from event-related potential (ERP) analyses have supported these claims, showing a stronger negative response in the brain signal soon after the presentation of congruent but not incongruent stimuli (Kovic et al., 2010). The authors suggest that the symbolic matching of auditory and visual stimuli enables a faster response owing to the stronger neural integration between these two representations.

Multimodal representations of speech have been widely discussed, independently of the sound symbolism literature. Language is accepted as a multimodal faculty, incorporating auditory and motoric representations (perception + production) as standard, and often in combination with visual mappings (i.e. visible articulatory gestures from another speaker; objects or images which correspond to the vocal representation). Furthermore, multimodality has been shown to be beneficial in language perception. An ERP study by Molholm and colleagues (2002) showed how multimodal experiences provide a processing advantage in adult speech perception: auditory-visual interactions generated 'surprisingly early' neural responses compared to the presentation of single-modality stimuli, suggesting an advantage for multisensory neural interactions in language processing. Westermann and Miranda's (2004) computational model of multimodal interactions in babbling highlights the importance of multimodal mapping in phonological development, as a tri-modal representation of speech activates mirror neurons in multiple domains during language learning. This relates empirically to an early experiment from Kuhl and Meltzoff (1982), where infants as young as 18 weeks were able to match auditory and visual stimuli, demonstrating that a multimodal representation of speech is already present in the first months of life.

1.3 Sound symbolism as a scaffold for language development

In their 'sound symbolism bootstrapping hypothesis', Imai and Kita (2014) posit that iconicity – incorporating both sound symbolism and onomatopoeia – is beneficial to early language development. They hypothesise that iconicity first leads infants to realise that sounds have meanings, and then facilitates the establishment of early word-meaning correspondences by prioritising the acquisition of those words which correspond iconically in the rich and varied adult input. Monaghan and colleagues (2014) make a similar suggestion with regard to the advantages of sound symbolism in language learning, stating that this feature may be "vitally important for language acquisition" (p.2). Using a corpus analysis of English words, they are able to support these claims from a production perspective by showing that words learned earlier in language development tend to be more systematically symbolic than those acquired later on. Their analysis can be criticized owing to the omission of data from infants under two years of age – indeed, one could argue that we cannot gain a true picture of acquisition if we begin analysis at this late stage of learning – though numerous other studies of younger infants appear to support their claims, with findings which reveal a sensitivity to systematic sound-meaning correspondences from a very young age. Both Ozturk and colleagues (2013) and Peña and colleagues (2011) tested four-month-old infants on their responses to matching and non-matching sound-symbol pairs (*bubu* to match a round shape and *kiki* an angular shape in Ozturk et al.'s experiment), and both studies found that infants were able to distinguish between congruent and incongruent pairs of stimuli. Ozturk and colleagues (2013) propose that these results demonstrate sensitivity to sound symbolism as an intrinsic feature of the cognitive system, since it is unlikely that these mappings derive from language exposure at such a young age. However, this leads to the question of why we find so few sound-meaning correspondences in language as a whole, if symbolic forms present a learning advantage in early development (Monaghan et al., 2012). In Monaghan and colleagues' (2011, 2012) discussion on this matter, the nature of common sound symbolic forms is considered, and it seems that these often relate to category-level features where there is often no need for precise differentiation between meanings: for example, *glisten*, *glimmer* and *gleam* share both phonological segments and semantic properties while *sheep* and *cow* share neither of the two – if these animals were instead labelled *feb* and *peb*, as in Monaghan and colleagues' (2011) example, a category-level distinction would be more problematic for the listener. An artificial language learning task showed this to be the case (Monaghan et al., 2011), as arbitrariness was found to be beneficial in communication through the maximisation of phonological information, thus leading to a more accurate establishment of meaning. However, this also depended heavily on context: a lack of contextual information had a negative effect on word learning.

Since Köhler's (1970) first inquiry into the human ability to make such sound-meaning correspondences, numerous studies have continued to explore the possibilities presented by sound symbolism. One study has shown that both Japanese adults and infants are able to make use of mimetic features in novel verb learning tasks (Imai et al., 2008), and an extension of the same task with English-speaking participants found the same responses across both adults and three-year-old infants with no experience of Japanese (Kantartzis et al., 2011). Most recently, an electroencephalography (EEG) task showed Japanese preverbal 11-month-olds to be sensitive to congruent auditory-visual mappings in novel words (*moma* to match a round shape, *kipi* to match a spiky shape; Asano et al., 2015), as multimodal correspondences were found to generate much quicker neurological responses in the matched condition than the mismatched condition. The authors suggest that an early advantage in the establishment of symbolic form-meaning mappings would enable infants to develop referential understanding, paving the way for the later acquisition of arbitrary word forms. However, the authors do not consider evidence from early word production to test whether this apparent processing advantage for non-arbitrary forms may facilitate language development outside the experimental setting.

1.4 Onomatopoeia and symbol formation

As far as lexical development is concerned, there is no clear evidence to show that infants prioritise the acquisition of non-arbitrary words in the earliest stages of language acquisition. That is, the 'frequency code' described by Ohala (1984) has not been found to apply in the early words that infants are reported to produce, at least not among infants under two years of age (but cf. Monaghan et al., 2014, as discussed above). However, none of the iconicity literature considers the extent to which infants produce onomatopoeia in their earliest words, and indeed, an observation of infants' early word forms appears to support the hypothesis that iconicity plays a role in the first stages of lexical development. Onomatopoeia are widely reported in infants' very earliest word forms: in Menn and Vihman's (2011) appendix of 48 infants' first five words, accounting for a range of 10 different languages, 20% were found to be onomatopoeic. Furthermore, Kern (2010) reported that words classed on the French adaptation of the MacArthur Communicative Development Inventory (CDI, Kern & Gayraud, 2010) as 'sound effects and animal sounds' were the most common in French infants' productive repertoires from eight months right up to 20 months old, and these forms contributed to over a third of the entire output of infants between the ages of eight and 16 months. Finally, Tardif and colleagues' (2008) cross-linguistic study shows a striking presence of onomatopoeia in early words, as on average the category of 'sound effects' was found to be second only to 'people' (e.g. *mummy, daddy, grandma*) in 970 infants' most common word categories across American English (29.5% of all words), Cantonese (40.6% of all words) and Putonghua (8.7% of all words).

Even in light of this evidence, there have been no empirical studies to-date which specifically consider onomatopoeia from this perspective, despite the wide array of experiments testing infants' sensitivity to sound symbolic pairings. Werner and Kaplan's (1963) account is the only in-depth consideration of onomatopoeia in child language development, yet this is exclusively theoretical. Indeed, their classic work *Symbol Formation* (1963) provides a comprehensive discussion of infants' "cognitive construction of the human world" (p.13), incorporating a detailed account of the importance of non-arbitrary sound-meaning links in the development of referential meaning. Possibly the earliest discussion of the role of onomatopoeia in infant language development was contributed by Farrar (1883), who described infants' production of onomatopoeia as a "natural instinct" (p.20), arguing that these forms play an important role as "stepping stones" to infant speech. Farrar argues determinedly in favour of Onomatopoeic Theory (Colman, 2009), which, similar to Cratylus' perspective discussed above, supposes that the imitation of sounds constituted the very first spoken words in human language. Farrar proposes that

> "the first men, in first exercising the faculty of speech, gave names to the animals around them, and...those names were onomatopoeic" (1883, p19).

While Onomatopoeic Theory has been widely dismissed – often condescendingly termed 'bow-wow theory' (Colman, 2009; Whitney, 1867) and indeed thought to be "wild and ridiculous" from Socrates' perspective (Jowett, 1953, 426b) – still Werner and Kaplan's account of infant language learning through the imitation of natural sounds has remained unquestioned in the wider development literature. Furthermore, although *Symbol Formation* is barely mentioned, the recent research into sound symbolism (e.g. Asano et al., 2015; Imai et al, 2008; Ozturk et al., 2013) only contributes to Werner and

Kaplan's perspective, with the assumption that sound-meaning correspondences may work as a bootstrapping mechanism in infant language learning (Imai & Kita, 2014). Werner and Kaplan relate to the undifferentiated state of the internal and external forms – that is, a referent and its symbol – in early development, as young infants are unable to grasp abstract conceptual representations. As a result, a form and its meaning are at first considered as one, as forms are derived from vocal imitations of natural sounds which eventually become meaningful protowords in the infant's early lexicon. The authors state that

> "[t]he widespread tendency of young children to imitate all kinds of noises – noises of objects as well as of persons (including those the children make themselves) – provides basic material *from which vocal depictive forms may be constructed*" (1963, p.101, italics added).

The authors detail the process of 'increasing differentiation', whereby an infant first uses an onomatopoeic form to refer to a referent, before moving towards a more abstract representation in the form of the conventional word as the need for a tangible 'likeness' between form and meaning decreases.

Werner and Kaplan's account of the ontogenesis of language from rudimentary vocal gestures corresponds to other emergentist accounts of language acquisition (for example, Bloom, 2000; Pierrehumbert, 2003; Thelen and Smith, 1994) which consider experience and self-organisation to be at the heart of cognitive development. The authors describe the unfolding of language from the most basic reflexes, as imitative vocal patterns emerge from communicative gestures such as grunts (McCune, 2008; McCune et al., 1996). Before these imitative vocal patterns are used for representation (termed 'designators' by the authors), the infant uses them as signals for indicating need. Eating, or the need for food, is a prime example of this signal-based vocal development, of which Werner and Kaplan cite numerous examples. The authors suggest that the /mː/ sound generated during periods of ingestion often becomes established by infants as a signal for requesting food, eventually forming part of a lexical unit in its own right to express food-related vocalisations, and, in some cases, leading to a more conventional vocal form such as *mother* /mama/. This example, which shows primordial vocalisations extending to become formalised meaningful units in an infant's vocabulary, is proposed as one of many such referential expansions that take place during an infant's vocal development. Such expansions are seen as a 'differentiation' process by Werner and Kaplan (1963), who consider the widening reference of an infant's vocal forms to reflect his growing understanding of himself as differentiated from his surrounding environment, and, most notably, his mother. As internalised vocal sounds become a signal of personal need, and later a name for an external object, meaning is seen to expand outwards from the infant's self. According to Werner and Kaplan's theory of symbol formation, this is a reflection of the infant's widening understanding of the world.

Werner and Kaplan regard onomatopoeic depiction as an essential part of the first step towards an infant's differentiation of himself as an individual within a wider world. They describe this as an ontogenetic phenomenon, as the relationship between both the inner meaning of a form and its external linguistic structure develop in tandem, interacting and modifying together as a result of experience of the world. As an infant moves from onomatopoeic to conventional word forms, one can often observe a transitory phase, whereby the infant combines the onomatopoeic form with the conventional adult form. This can be seen in the early words of numerous infants, including Hildegard Leopold (Leopold, 1939), who transitions from producing train as /dʒudʒu/ or /tʃutʃu/ (*choo choo*, 1;8) to /dʒudʒute/*choo choo train* at 1;11. Werner and Kaplan attribute this phase to a gradual shift in representation, during which time the infant is able to 'mould' the new word form to her internal idea of the object in question. From here, "the new vocal material becomes dynamically-physiognomically organized to fit the event represented" and now portrays the object "not materially, but through the way in which is it apprehended internally by the user" (p.129). This change in representation takes place throughout the infant's lexicon, as the symbolic connectedness between vocalisations and their corresponding objects develops towards a more abstract relationship, considered by some to be one of external association. Others, such as Werner and Kaplan themselves, consider this to be a 'physiognomization' of linguistic forms; that is, forms which are internally linked to, though externally remote from, their referents (Werner & Kaplan, 1963). This view extends beyond onomatopoeia to include infants' expressive representations of objects as a whole, and while it can be considered as being a rather radical account of the unfolding of language, Werner and Kaplan's (1963) perspectives on form-meaning correspondences are not too far removed from those of the more recent sound symbolism literature. Indeed, Asano and colleagues (2015) also posit that the early bias for symbolic forms supports development towards referential understanding in general, while Imai and Kita (2014, 11) propose that "pre-verbal infants detect sound symbolism in unfamiliar words and process them as if they were real words", thus supporting the establishment of sound-meaning mappings in early development. These points support a role for iconic forms in early language development in terms of a bootstrapping mechanism, which promotes early word learning and also sets the stage for the acquisition of arbitrary forms in later development.

1.5 Iconicity in sign and gesture

Iconicity in language does not exist exclusively in speech: a rich inventory of symbol-referent congruencies occurs in signed languages, much more so than can be typically found in spoken language. In their comprehensive review of the sign language acquisition literature, Meier and Newport (1990) note that this provides an opportunity for the comparison of iconicity in language learning. However, despite the increased presence of iconicity in signed languages, no iconic advantage has been found in sign language acquisition, which Meier and Newport (1990) acknowledge as being somewhat surprising. Orlansky and Bonvillian (1984) found that infants acquiring sign language showed no preference for iconic over non-iconic signs: iconicity was present in around 30% of the first 10 signs acquired by deaf infants, increasing to one third by 18 months. Although sign language lexicons possess an increased number of iconic forms relative to spoken languages, the proportions reported in Orlansky and Bonvillian's (1984) study represent typical findings in terms of hearing infants' early production of onomatopoeia, as also found in studies by Kern (2010) and Tardiff and colleagues (2008). Meier and colleagues (2008) found no advantage for iconicity in sign acquisition either, as deaf infants did not draw upon iconicity as a feature of the learning process: the authors hypothesised that infants would express enhanced iconicity in the early production of signs, but this was found not to be the case, and in fact more forms were found to be produced with fewer iconic features.

The literature on gesture presents a more confusing picture, as evidence of its advantage in the acquisition of speech is mixed. Namy, Campbell and Tomasello's (2004) study observing hearing infants' acquisition of iconic versus arbitrary gestures reported similar results to those of Orlansky and Bonvillian (1984) and Meier and colleagues (2008): no learning advantage was found for iconic gestures across infants of 18 and 26 months and four years of age. However, Acredolo and Goodwyn (1988) found that infants' early use of symbolic gestures in reference to objects (for example, the use of a 'panting' gesture to designate *dog* or spreading of the arms to represent *aeroplane*) correlated with their lexical development, leading them to reach the 10-word point earlier than those infants who did not use many symbolic object gestures. Similarly, Rowe and Goldin-Meadow (2009) found that, for older infants of 18 months, the number of different gestures used to express different meanings (e.g. pointing to an object in reference to it, head-shaking to represent *no*, arms spread to depict *aeroplane*) was indicative of the infants' vocabulary at 42 months: infants who used a larger variety of gestures at 18 months had larger vocabularies two years later. In the input, caregivers' use of gesture in combination with spoken language has also been shown to have a strong positive effect on lexical acquisition for both typically developing infants (Goodwyn, Acredolo & Brown, 2000) and infants with specific language impairment (Lüke, 2015; Vogt & Kauschke, 2015). Finally, a study by Lieberth and Gamble (1991) tested adults with normal hearing on their ability to learn and retain iconic versus arbitrary signs. No difference was found in the adults' ability to retain the two sets of signs in a short-term memory task, although a decrease was observed in their ability to retain arbitrary, but not iconic, signs over time.

An understanding of how iconicity in signs and gestures may support language learning can move us closer towards the formulation of a plausible hypothesis regarding the nature of iconic form-meaning relationships in spoken language. However, there appears to be a divide in this evidence, as we find both a case against a role for iconicity (based on the evidence from sign language) and a case for some role for iconicity (based on the gesture literature). This confounding evidence is difficult to interpret: on the one hand, it supports Werner and Kaplan's (1963) claims regarding the undifferentiatied state of object and symbol (in this case, a referent and its related iconic gesture) in early development, and the process of increasing differentiation as the 'distance' between a form and its meaning increases over time. Acredolo and Goodwyn (1988) posit that action-based iconic symbols may have an increased advantage in language learning owing to the less demanding memory load that is provided through visual access to the referent, which can also serve as a transitional state prior to the acquisition of the vocal form. On the other hand, however, the sign language literature shows how iconic forms do not serve any learning advantage in the absence of spoken language, which suggests that it is the pairing of speech and gesture that may corroborate an infant's learning of particular word forms. This returns us to the question of multimodality in language acquisition (e.g. Ramachandran & Hubbard, 2001), and leads us to posit that it is not the motivated form-meaning correspondences of iconic gestures that are beneficial to the infant, but rather the combination of sensorimotor cues which may facilitate processing and provide an early mnemonic advantage.

With signs and gestures, as with speech, the nature of iconicity has been called into question. Slobin and colleagues (2003) consider the nature of iconic signs that bear visual resemblance to their referent, and note how the iconicity-arbitrariness debate is not as dichotomous as often appears: the authors propose that this should instead be considered as a spectrum, whereby even the most conventionalised signs can maintain some visual resemblance to their referent. This point of view can be extended to the consideration of onomatopoeia in spoken language, where the question of iconicity versus arbitrariness is not always as clear as might immediately be assumed. As will be discussed below, the nature of these arguments only adds to the uncertainty that surrounds the question of how to approach these forms in linguistic analyses.

1.6 Onomatopoeia in the development literature

The literature on onomatopoeia in language development is limited, especially with regard to phonological development. Werner and Kaplan (1963) and Farrar (1883) provide the only thorough discussions offering explanations as to why infants might produce so many onomatopoeia in their earliest words. Although not wholly identical, these two perspectives present a largely one-sided account of onomatopoeia in infant word production, as both posit that the less-than-arbitrary sound-meaning correspondences of these forms make them intrinsically learnable. De Saussure (1962/15), on the other hand, claims that onomatopoeia are largely conventionalised. Indeed, he goes as far as stating that these forms are in fact arbitrary “to a certain extent” (p.102, own translation), describing an onomatopoeic word as being “only an approximate imitation and already partly conventionalised” (p.102, own translation). He goes on to argue that words which may in their essence be onomatopoeic lose their symbolic character over time, taking on an arbitrary linguistic form which is devoid of all iconicity. Furthermore, he rebuts arguments for phonological sound symbolism in one short paragraph, in which he explores the etymology of certain supposedly symbolic forms (French *fouet* ‘whip’, for example, which derives from the Latin *fāgus* ‘beech tree’) and denounces any symbolic correspondences as being “a fortuitous result of phonetic evolution” (p.102, own translation). A brief consideration of infants’ use of onomatopoeia is enough to dispute claims suggesting that these forms are anything other than conventionalised. Many infants growing up in urban environments may produce forms such as *moo* and *baa* in reference to cows and sheep without ever having heard these animals make any such sounds, and for the most part trains can no longer be realistically considered to make a sound resembling *choo choo*. Indeed, we must question whether infants' representations of these so-called onomatopoeic forms is any different from that of the words *ball* or *banana*, for instance, which are accepted as being learned from the adult input in reference to the objects in question. However, it cannot be denied that infants do tend to imitate sounds, creating their own idiosyncratic onomatopoeia, which are often reported in the literature to be used in reference to specific objects or phenomena. For example, Elsen's (1991) daughter Annalena produces her own imitative forms to refer to numerous animals, reported by the author to be created by the infant based on her own interpretation of the sounds that these animals make. Evidence can also be found for the role of the input in infants' production of onomatopoeia: Leopold (1939) reports how his daughter Hildegard[1] learned the sound "sch, sch, sch!" (p.121) from her grandfather, used in games relating to trains. Confounding evidence has therefore been presented with regard to infants' acquisition of onomatopoeic forms: while it seems that infants do indeed create their own referential words based on onomatopoeia-style imitations, we also find many examples of onomatopoeic forms which are learned from the adult input. Here we must ask what status these two differing forms of onomatopoeia may take in infant language development, and indeed whether either or both may directly facilitate word learning.

With regard to the proposed learning advantage for onomatopoeia, the sound symbolism literature – as reviewed in Imai and Kita's sound symbolism bootstrapping hypothesis (2014) – is the only source of empirical evidence that may support these claims. Some other suggestions have been put forward as researchers attempt to deal with the puzzling nature of onomatopoeia in early infant production, but no attempt has been made to further pursue the question of why these forms are so abundant in early speech. In a study of syllabification in Finnish infants' language development, Kunnari (2002) debates how best to deal with the large number of onomatopoeic words in her dataset, which appear to be treated differently to non-onomatopoeic words in the early output. In her analysis, she initially considers these forms as part of the wider results, but also analyses them separately from the other words in her study, dealing

[1] The language development of both of the infants discussed here will be considered later in this thesis.

with the effect that onomatopoeia has specifically on Finnish infants' syllabification. Onomatopoeia are found to alter the infants' data early on in word production, as these forms are often monosyllabic in Finnish, contrasting with the non-onomatopoeic forms which were largely polysyllabic. This affected the results up until the 25-word point (the point at which the infants produced 25 words or more in a single recording session): at the 15-word point, monosyllabic onomatopoeia constituted 24% of the infants' words in her analysis. Kunnari discusses why onomatopoeia may be so present in her data, and why they appear to differ from the common production patterns that were identified in non-onomatopoeic words. She observes that even the polysyllabic onomatopoeia are produced more accurately than other words with the same number of syllables, suggesting an articulatory advantage for these forms in early production. Kunnari proposes that onomatopoeia should be treated separately from non-onomatopoeia in similar studies of language development, since these forms constitute such a large proportion of the early output, and appear to be more accurately produced than might be expected. It is also suggested that this may be due to these forms being pragmatically or prosodically more salient, and that they have "quite an easily mastered articulatory shape" (p.133). Indeed, when we consider the phonological nature of common onomatopoeic forms in English we see that this is the case: reduplication and simple CV syllables are common in these words (for example, *woof woof*, *moo*), and this is typically the case across languages (Dutch: *woef woef*, *moe*, Hebrew: *how how, moo*, Greek: *gav gav, moo*; Abbot, 2004). This corresponds to the common phonological features of infant speech, as simple syllable structures and reduplication are found to be common features of infant data across languages (Ferguson, 1983; Vihman, 2015). Kunnari (2002) does not consider any form of bootstrapping role for onomatopoeia, however, going against the general trend which assumes that onomatopoeia play a special role in language learning. Savinainen-Makkonen (1998, 2007) also comments on the presence of onomatopoeia in Finnish infants' early outputs, observing that these forms were often produced more accurately than non-onomatopoeic targets with the same number of syllables (1998, cited in Kunnari, 2002). The author also observes that onomatopoeic forms often have a segmental structure that is not typical of Finnish (Savinainen-Makkonen, 2007), including word-final consonants and consonant clusters which are rarely found in the target language.

Jakobson (1980; original publication 1941) discusses the potential importance of onomatopoeia in establishing early phonological capacities. He begins by highlighting the problems that infants encounter in this task, whereby, in order to develop a stable and usable phonological system, infants must be able to store segments in their memory as well as recognising and reproducing them whenever necessary. He refers to this as the stabilisation of "arbitrary sound distinctions aimed at meaning" (p.25), and thus acknowledges the difficult task that infants face in building up an arbitrary linguistic system from scratch when there are no meaningful correspondences between a sound and its referent. He goes on to discuss the role of onomatopoeia in this process, but rather than suggesting that the symbolic link between sound and meaning in onomatopoeia may provide a scaffolding opportunity in development, instead he discusses the role that these forms may play in 'linguistic training' (p.27). Citing numerous examples from the literature, Jakobson (1941/80) suggests that the production of onomatopoeic sound imitations in early word production may facilitate phonological development through the provision of production practice. An infant may produce the segment /ɹ/ when imitating a car, for example, which provides an opportunity to practice this phoneme without the challenges provided by the surrounding segments that constitute a whole word. Indeed, even when certain segments are too difficult for an infant to produce in the context of a word, they can still be produced reliably in onomatopoeic imitation of a sound. Jakobson relates to this as "the expressive value of the extraordinary rather than the desire for faithful sound imitation" (1980, p.26), as the unprescribed structure of an imitative onomatopoeic form gives an infant the freedom to produce segments that are not possible when constrained by surrounding articulatory demands. Jakobson claims that this facilitates the stabilisation of these phonemes, thus contributing to the establishment of the phonological system.

Despite the various suggestions as to the nature of onomatopoeia in early word production, no empirical research has been conducted to determine precisely what significance (if any) these forms might have on infants in the process of establishing the rudiments of their lexicon. Vihman (2014) concludes that "the question of just how these less than fully arbitrary transitional forms support early word learning has seldom been addressed" (p.161), but here we must question to what extent onomatopoeia have been addressed at all in the field of phonological development. Indeed, many studies can be found in the literature which choose to omit onomatopoeia from the dataset altogether: Behrens (2006) views these forms as "nonce words" (p.15), disregarding them alongside other 'meaningless' aspects of speech such as "hesitation markers…and noninterpretable words" (p.15), and similarly Fikkert and Levelt (2008) do not consider onomatopoeia in their consideration of place of articulation in phonological development. Since 2010, over one third of articles published in *Journal of Child Language* which consider infants' early perception or production of speech have explicitly omitted onomatopoeia from the analysis, while a further 44% fail to specify whether or not onomatopoeia are included in the dataset. In part this is likely to be due to the marginal status of onomatopoeia in the adult language, leading these forms to be characterised as such in infant speech as well. However, while Kunnari (2002) has shown these forms to present a problematic question to consider in the analysis of infant speech, the extent to which they are present in many infants' early outputs shows them to be more than simply marginal to child language development. Indeed, in viewing the trajectory from first words to adult speech as a linear passage, with the adult model as the only valuable perspective from which to consider development, we may omit essential evidence regarding the nature of language acquisition.

This thesis will take into account the use of onomatopoeia in both infant and caregiver speech. While we assume that many of these forms are learned from the conventionalised onomatopoeia that are presented in the input, we leave room for the consideration of words which the infant (or caregiver) creates, as reported in accounts from Leopold (1939), amongst others. Here it is hypothesised that these words do not have an intrinsically-determined meaning that facilitates early language development, but instead that they are acquired like any other aspect of language: as a result of exposure in the input and experience in the output. However, we also propose that onomatopoeia may play a unique role in lexical development, as these forms stand on the peripheries of language and are thus exceptional in terms of both the input and the output.

First, we acknowledge that many onomatopoeia constitute simple structures such as reduplicated syllables (for example, *woof woof*, *quack quack*, Ferguson, 1983) and CV structures such as *moo* and *baa*, as posited in the literature to be an advantage of onomatopoeia (Kunnari, 2002). On the other hand, the idiosyncratic nature of onomatopoeia provides a certain level of phonological under-specification, which not only allows for a more flexible production of these forms on the infant's part, but also enables the production practice of individual segments, as referred to in Jakobson's (1941/80) analysis. Here we propose that *the phonological features common to onomatopoeia* *make them more producible in the early output* through being phonologically more flexible as well as prosodically easier for infants to produce. Second, as Kunnari (2002) suggests, these forms may be particularly salient in the input owing to the use of idiosyncratic prosodic features to imitate the real-world sounds that the onomatopoeic word is assumed to represent. This also corresponds to their status as lexical features of infant-directed speech (Ferguson, 1983). We therefore hypothesise that *onomatopoeia will be particularly distinctive in the early input* by virtue of their marginal status – that of 'sound effect words' – in the adult language. Finally, we posit that infants' use of onomatopoeia is primarily driven by the caregiver input, as suggested in Leopold's (1939) account, as well as many other studies of language acquisition (Kauschke and colleagues, 2002, 2007). Thus, *so long as infants are exposed to these forms in the caregiver's speech*, our overall hypothesis proposes that *phonological advantages presented by onomatopoeia in the input as well as the output will bring about their common production in infants' earliest words*. This will be tested through a consideration of onomatopoeia from three distinct perspectives:

1) *Onomatopoeia in production*. How do onomatopoeia fit within infants' general phonological development, and is there a phonological advantage for these forms in the early output?
2) *Onomatopoeia in perception*. How are these forms presented to infants in the early input, and do they fit in with or stand out from other, non-onomatopoeic forms?
3) *Onomatopoeia in interactions*. What role do these forms play in early infant-caregiver interactions, and how does this facilitate infants' representation of onomatopoeia in relation to their corresponding non-onomatopoeic word (e.g. *woof* versus *dog*)?

1.7 Structure of this thesis

In order to explore this topic thoroughly, early production and perception will both be considered. First we will analyse any general trends in the acquisition of onomatopoeia in early language development from the earliest words of nine infants acquiring a range of six different languages (Chapter 2). This will provide a perspective on how onomatopoeia are acquired over time, and how this fits into the wider trajectory of lexical development, with an analysis of these forms in the infants' early phonological development. Following this, Chapter 3 will provide a more in-depth view of onomatopoeia in early language production through a case study of one infant acquiring

German. Here we will continue to consider onomatopoeia in phonological development, with an analysis of the specific segmental properties of these forms, and how they fit within the infant's wider phonological inventory. Chapters 4 and 5 present perspectives from early language perception through two eye-tracking studies which question the proposed advantage for these forms in terms of their symbol-referent correspondences. Here we will test infants' on-line processing of onomatopoeia in order to determine whether there is any perceptual advantage for these forms in early development, either in the facilitated mapping of forms to meanings, or in the presentation of these words with appropriate 'sound effects' which may make them stand out from the input. The last two chapters will then consider onomatopoeia in the caregiver input, first with an analysis of onomatopoeia in infant-directed speech (IDS), comparing the production of these forms with equivalent non-onomatopoeic words in recordings of twelve caregiver-infant interactions (Chapter 6). Finally, Chapter 7 presents longitudinal video data from eight caregiver-infant dyads, where we will observe the changing nature of onomatopoeic production over time, while also considering the specific use of these forms across individual dyads. Together these six analyses will provide a comprehensive view of onomatopoeia in early language production, taking both perception and production into account as well as the interaction of the two. Furthermore, longitudinal analyses will provide a perspective that has not been considered in the literature, enabling us to determine how an infant's language ability at any one point governs the production of onomatopoeia. This unique combination of perspectives will address the questions raised in the literature regarding the symbolic properties of these forms, as well as the suggestions made by Kunnari (2002) and Jakobson (1941/80) with reference to the role of onomatopoeia in phonological development. Finally, we will attempt to determine the extent to which these forms are conventionalised, thus returning to the arguments presented by de Saussure (1915/1962) as well as Plato (Jowett, 1953) regarding arbitrariness in language.

1.8 Terminology

Here onomatopoeic words (OWs) are defined as words which attempt to imitate a sound from the environment, whether or not these are conventional forms such as *woof woof* and *bang*, or idiosyncratic forms which have been created by the infant or the caregiver based on their own interpretation of a sound. In many cases these forms may be used as referents – for example, a dog may be referred to as a *woof woof* – and in these cases the OW form will have an equivalent form in the adult language. In such cases, where an OW is used as a referent for an object or phenomenon, the typical adult form which the infant will eventually use referentially will often be considered in conjunction with the OW, and will be referred to as the 'conventional word' (CW). However, in the case of some OWs, such as *boom*, *bang* and *pop*, no equivalent CW form exists (though this is to some extent idiosyncratic, as Werner and Kaplan (1963) cite one example of an infant referring to a hammer as a *boom*, for example). In terms of these OW-CW dyads, the CW form will be considered as the target, or gloss, and will appear in small caps: DOG therefore relates to the OW form *woof woof* and the CW form *dog*.

1.9 Statistical analyses

The analyses in this thesis will mainly be carried out using linear mixed-effects modelling. This method is becoming increasingly popular in psycholinguistics (Barr et al., 2013), as it allows researchers to combine fixed and random effects within the same model while also taking into account the random differences across subjects in an experiment (Winter, 2013).

The majority of data for analysis in this thesis combines fixed and random effects. Furthermore, in some cases a number of different subjects are tested on a number of different items; mixed-effects models allow us to combine multiple-subject responses to multiple items, while also accounting for the between-subject error that we can expect in reaction time or voice pitch across participants, for example. In experiments with multiple items per participant (as in Chapters 4, 5 and 6), ANOVAs would require the combination of two statistical models, one accounting for by-subject data and the other for by-item data. With mixed-effects, we are able to add all data to a single model, which allows us to account for all error within that one model in the form of by-subject and by-item random slopes (see Barr et al., 2013, for a review), thus increasing the power of the analysis. Random slopes take into account any variability across subjects in relation to items (Winter, 2013): if a participant has a particular response to an item that is not consistent across subjects, random slopes can account for this. For example, if an infant has a particular interest in cows, and this is not consistent across all participants, random slopes take into account the individual infant's longer response in the COW trials relative to the rest of the data. The inclusion of random slopes is a much-discussed aspect of linear mixed-effects analysis, but as Barr and colleagues (2013) conclude in their extensive discussion on this topic, the use of 'maximal' models (i.e. including random slopes as standard) is necessary for confirmatory hypothesis testing if we are to take a conservative approach to our data.

In this thesis, linear mixed-effects models will be generated in R (R core team, 2014), using the *lmer()* function in the lme4 package (Bates, Maechler & Bolker, 2012). Unless otherwise specified, by-subject and (where appropriate) by-item random slopes will be included in all models. P values will be obtained using likelihood ratios to compare the full model with the effect in question against the model without the effect in question. In cases where further exploratory analysis is required, post-hoc ANOVAs and t-tests will follow up the results. All reported t-tests will be two-tailed.

SECTION I

Onomatopoeia and templates in phonological development

In this first section, a template-based phonological analysis will be carried out in consideration of infants' production of onomatopoeia. This will allow us to identify how these forms may match, or indeed contrast with, the typical trajectory of the early output, providing a perspective on the role that onomatopoeia play in early word production. It has been suggested that onomatopoeia stand out from infants' general output forms both phonetically (Savinainen-Makkonen, 2007) and prosodically (Kunnari, 2002), and that the phonologically simple and under-specified nature of OWs enables a more flexible approach to their production (Jakobson, 1941/80). A template-based analysis of infants' outputs will allow us to consider their OW production in line with these interpretations, assessing how these forms fit into the infants' wider phonological development.

In the 1970s a number of studies of infant language development brought to light evidence of 'whole-word phonology' (Ferguson & Farwell, 1975; Macken, 1979; Waterson, 1971). These studies noted production patterns in infants' early words which appeared to be individual to the infant in question, referred to by Ferguson and Farwell (1975) as the systematization of the early output. In their consideration of three infants' phonological development, the authors consider the 'individual strategies' that these infants take in the organisation of the earliest phonological system, which includes

> "preferences for certain sounds, sound classes, or features ('favorite sounds'); extensive use of reduplication…; preferences for either lexical expansion or phonological differentiation at the expense of the

> other; and persistent avoidance of particular 'problem sounds'" (1975, p.436).

This approach reflected an important shift away from Jakobson's (1941/80) quest to identify phonological universals in child language, which is discussed in detail in Ferguson and Farwell's account, and an increasing body of research has since been dedicated to the identification of these systematic phonological patterns, or 'templates' (Vihman, 2014), in early language production. Longitudinal accounts of infant language development, dating from as far back as the early 20th century, have been retrospectively considered within this framework (Vihman & Keren-Portnoy, 2013; Vihman, 2015), while new data has broadened the understanding of phonological templates through the analysis of infants acquiring different languages (Vihman, 2010) and multiple languages (Vihman, 2015), twins (Smith, 2011) and 'late talkers' (Vihman, DePaolis & Keren-Portnoy, 2009; Vihman et al., 2013).

Vihman (2014) discusses templates and whole-word phonology in her chapter on 'Functionalist or Emergentist Models', in line with the usage-based theories presented by Bybee (2001) and Pierrehumbert (2003), among others. The whole-word approach considers even the earliest vocal articulations as being central to language learning, positing babbling as the first stage in early word production, in contrast with Jakobson's (1941/80) belief that babbling serves no purpose in early communicative development. As detailed by Waterson (1971), whole-word phonology assumes an independent phonological system in early word production which is still wholly related to the adult language presented in the input. In an analysis of her son's language acquisition, Waterson describes the early output as a set of 'schemata' (p.206), which are shaped by the infants' individual impressions of words as gestalts – that is, *whole words* – and are perceived by the infant from the utterances heard frequently in the input. Initially, the infant is considered to be without a phonological system, which is then built through the perception and reproduction of the most frequently-encountered phonological structures. Through this process, the infant builds up his 'own system' (p. 205), whereby one can relate consistent patterns from the production of his individual words to the segments and words perceived in the adult language. Waterson's lack of clarity regarding the nature of the input and its salient features has been highlighted as a gap in her analysis (Vihman, 2014; Vihman & Croft, 2007), but a similar account published by

Menn in the same year draws further upon Waterson's approach, and presents the argument that 'the word is an entity, stored and accessed as a block' (1971, p.247).

The whole-word approach to language production, as described above, is re-interpreted as a template-based approach by Vihman and Croft (2007), who propose that the early production of whole-word shapes brings about the establishment of phonological templates, constituting phonological features and prosodic structures that are typical of the ambient language while also being motorically accessible for infants in the first stages of word-use. An infant's templates are often derived from the preferred babbling patterns of the pre-linguistic period, constituting simple prosodic shapes and a small number of well-established consonants, or 'vocal motor schemes' (VMS; Vihman & McCune, 2001). These early patterns may then go on to form templates, which function as a systematic response to the articulatory challenges presented to the infant by the ambient language. As the infant gains further linguistic experience, templates are replaced with more complex prosodic structures and phonological categories.

Infants' earliest words are often found to be accurately matched to the adult form through the item learning of words which automatically fit the infant's production capacity. Examples can also be found of 'progressive idioms' (Ferguson & Farwell, 1975), whereby the infant produces a word with unexpected accuracy, presenting a mismatch in terms of their wider phonological output. The majority of early accurate forms often match the established prosodic patterns that the infant has been producing since the babbling phase, creating a systematically-acquired lexicon that is largely based on the infant's preferred and well-rehearsed output patterns. The steady accumulation of new words to fit these patterns – known as 'selected' words – enables the construction of the early phonological system. Over time, these patterns are overgeneralized through the production of words which do not automatically match the well-rehearsed prosodic structures, causing words to be 'adapted' in line with the ever-increasing systematicity of the early output. Vihman (2015) accounts for this phenomenon as a mismatch between the infant's 'ambition' to produce increasingly complex word forms and his phonological limitations in terms of articulatory control, planning and memory. Vihman specifies the nature of templates through the identification of either the 'overuse' of particular prosodic structures, or the systematic adaptation of target words to match the specific features of the particular template. An example is given from Priestley (1977), whose son implemented the template <CVjVC>

in the production of disyllabic words with codas, as in *farmer* [fajam], *tiger* [tajak] and *basket* [bajak]: here we can see the infant's wholesale approach to a specific whole-word challenge, dealt with while remaining within the confines of a limited phonological system. Over time, and depending on the individual infant's approach to lexical acquisition, infants can be seen to move towards a more adult-like system for word production: word-based adaptations shift towards phonological substitutions, and thus from a whole-word to a segment-based approach to language learning.

From this perspective we can clearly see how templates might affect word learning, through facilitating lexical development within the infant's existing phonological capacity while also overcoming issues of memory (see Keren-Portnoy et al., 2010) and planning. Furthermore, any approach which promotes word production also provides the infant with essential articulatory refinement through production practice, thus bringing about further phonological development over time. Naturally, however, the adaptation of words to fit an infant's templates will lead to a loss of accuracy across word forms both old and new: while newly-acquired words are automatically adapted to fit the infant's systematic templates, so too are previously accurate words adapted, leading to a regression in the overall accuracy of the infant's output. This 'U-shaped curve' seen in accuracy over time can thus be attributed to the infant's increasingly systematic approach to word production, followed by a gradual shift back towards a more adult-like phonological model.

While templates are not universal across either infants or languages, Vihman (2010, p.278) reports an "overall cross-linguistic similarity in output forms", while features such as reduplication, consonant harmony, truncation and cluster reduction are described by Ingram (1974) as common phonological processes in the early output. However, in contrast to these common tendencies, the idiosyncratic nature of infants' templates has also been widely established, both in terms of the language being acquired, whereby the particular challenges of the target language are dealt with using the same phonological processes, and in terms of the highly idiosyncratic approaches adopted by infants to contend with the individual problems faced in the output (Vihman, 2010; Vihman & Keren-Portnoy, 2013).

The trajectory from babbling to templates, however, is not a sequential or exact process, and while the literature reports many cases of highly systematic template-use among infants (e.g. Macken, 1979; Priestley, 1977; Waterson, 1971), there are many more infants whose early output cannot be easily identified as containing templates, or whose template-use is highly variable and thus difficult to determine. As Vihman (2010; 2015) has acknowledged, when accounting for the use of different prosodic shapes in the early stages of word production, it is common to find the same target produced with a varied set of prosodic structures. For example, Vihman's (2015) son Raivo, acquiring English and Estonian, is reported to produce *viska* 'throw' as [is], [iɬ] and [ɬ̩], incorporating two templates from his early output here: <(C)VC> and <C>. Similarly, some words may not fit any of the infant's common prosodic shapes, as in Pačesová's (1968) son P, acquiring Czech, who produces some anomalously complex word shapes in his first 100 words which do not fit his otherwise systematic early output. These include *maso* 'meat' [maso] (<C1VC2V>), *utíkat* 'run' [tika:m] (<C1VC2VC3>) and *hop* 'hop' [hap] (<C1VC2>). Further template-use in P's output will be discussed later in this section.

Another problem encountered in the analysis of an infant's early templates is the contrast between the terminologies *prosodic structure* and *template*. While the two phenomena are not the same, the interaction of the infant's common prosodic structures and established templates is an important feature in the early output, and will feature prominently in the two chapters which follow. As early as the babbling phase, infants produce prosodic structures which apply consonant and vowel features in a range of formulations, thus shaping individual preferences in terms of both the segmental and the phonological properties of the early output. The characteristics of the earliest lexical items 'match' an infant's prosodic structures, as established through the vocal practice of babbling (Vihman et al., 2009); these words are often accurate reproductions of the target word, and are thus considered to be SELECTED for that particular prosodic structure. When the infant faces the problem of a new word which does not match one of these structures, that word is ADAPTED to match the rest of the output and the phonological capacity that is available to him at that stage. When a particular prosodic structure is used to select a disproportionate number of words in the early output, or when a disproportionate number of words are adapted to fit that structure, we see evidence for the use of a phonological template. This reflects the use of a well-rehearsed phonological structure to contend with challenges in the output, and thus the systematization of the early output begins.

In this sense, a template can be seen as an extension of the common prosodic structures found in an infant's output, whereby a certain pattern is overgeneralised to bring about vocabulary growth. In this section we will consider the use of templates in the early output, first analysing nine infants' first 100 word forms for the use of common prosodic structures which are systematically applied across the developing vocabulary to become templates. This analysis will focus on the acquisition of OWs during the early stages of word production, and how these forms relate to each infant's individual use of templates.

We will then focus on one of these infants' further lexical and phonological development up until the 500-word point, by which time the systematic implementation of templates will be well-established, making it easier to observe any specific function for OWs within a wider linguistic framework. These two analyses of the early output will introduce the reality of OWs in the early vocabulary and set the scene for a broader analysis of the presentation and perception of these words in early development.

2. Diary analysis I: Trends and templates in OW acquisition

2.1 Introduction

In this chapter we will consider data from nine infants in order to understand the early stages of language development in relation to the acquisition of OWs. Here we raise the question of whether these forms deviate phonologically from the infants' developing system, or whether they fit systematically within the early lexicon. Kunnari (2002) refers to the simple articulatory structures of OWs, which she posits as a reason for their abundance in early production; here we will test this proposal through the analysis of infants' first 100 words, adopting Vihman's (2015) analysis of infants' prosodic structures to compare OW production with that of the wider lexicon.

This overview will first provide a broad perspective on OW acquisition in early word learning, followed by a close-up phonological analysis which will allow us to identify some specific and perhaps surprising patterns across both the infants' production of OWs and the phonological properties of these forms. An analysis of the early outputs of nine infants in terms of their lexical and phonological development, across word production in general and OWs more specifically, will enable us to determine whether the fleeting presence of OWs in the early lexicon – a prominent aspect of the early vocabulary yet only marginal in the adult language – has any facilitative, innovative or functional role in language development.

The nine infants come from a range of language backgrounds: between them they are acquiring six different languages, and four of the infants are raised bilingually, all with English as one of their languages. All of the datasets are diary recordings, of which eight were compiled by parent linguists, and are sufficiently detailed to provide a clear and thorough insight into the infants' language learning.

2.2 Methodology

2.2.1 Data

This study aimed to capture the very first 100 words produced by each infant, and data sources were limited accordingly: data was ruled out if recording had begun after the onset of word production, and any sources which did not provide a regular account of an infant's word learning had to be disregarded. The nature of diary accounts makes it impossible to ensure direct parallels across the eight datasets analysed, and the extent of detail provided by the researchers differs for the infants analysed here. However, the infants' first 100 words were recorded in all datasets as a minimum, with a (usually broad) phonetic transcription provided for each. Two of the infants (Maarja and Kaia) are sisters, and so here we can assume similarities in data collection methods.

The datasets do not always specify whether the infants' words were produced spontaneously or as imitations of adult speech, and so unless reports suggest otherwise all word inventories were taken as accurate depictions of the infants' developing lexicons. However, forms which are reported in the datasets as words but which correspond more accurately to exclamations or vocal gestures, such as the 'demonstrative interjection' [ʔəʔ] found in Hildegard's data (0;8, Leopold, 1939) and the use of [aː] to represent 'I'm happy' (M, Deuchar & Quay, 2000, p. 119) were excluded from the analysis.

2.2.2 Participants

Data were collected from longitudinal diary studies of infants acquiring a range of languages (see Table 2.1). In all cases word production was recorded in a diary format: eight of the nine datasets were parent linguists' diary notations, taken while observing their infant's language development for their own research purposes. These were all hand-written accounts, and some were supported by audio back-up, though these recordings were not considered in this analysis as they were accessible for only one of the datasets (Laura: Braunwald, 1976). Trevor's data was gathered by his own parents, and in this case they were not linguists, but participants in a wider project (Compton & Streeter, 1977; Pater, 1997) where speech pathologists were recruited and trained in phonetic transcription in order to record their own infant's speech. Data was hand-written in notebooks and at least four hours of recording per week was provided throughout the data collection period.

Table 2.1: Sources of data used in this analysis.

Child	Sex	Language	Age	Data source
Annalena	F	German	0;8-1;1	Elsen (1991)
Hildegard	F	German & English (US)	0;9-1;7	Leopold (1939)
Kaia	F	Estonian & English (US)	0;11-1;8	M. M. Vihman (unpublished data)
Keren	F	Hebrew	0;10-1;3	Dromi (1987)
Laura	F	English (US)	1;3-1;4	Braunwald (1976)
Maarja	F	Estonian & English (US)	1;0-1;5	Vihman & Vihman (2011)
M	F	Spanish & English (UK)	0;10-1;6	Deuchar & Quay (2000)
P	M	Czech	0;10-1;5	Pačesová (1968)
Trevor*	M	English (US)	0;8-1;3	Compton & Streeter (1977), Pater (1997)

* Data was collected by parents who were not the researchers.

2.3 General analysis

Data was analysed from each infant's very first word, and considered until the infant had a total of 100 different word forms in their lexicon. This allows a like-for-like comparison of vocabulary development, despite the fact that the infants differ in age: Annalena and Trevor were the first to begin speaking, both at 0;8, while Laura was the last, at 1;3. Kaia is reported to produce her first word – *kiisu/kitty*, produced in a whisper as [kiːtɔ] – at 0;11, but then we observe a long delay of over three months before she produces her second word *nämma* 'yum' [mæmː] at 1;2.

For the four bilingual infants, words which differ in the two languages and were acquired separately (for example, *granny* and *abuela* 'grandma' in M's data) were considered as two separate word forms, while words which are phonetically similar across the two languages such that the specific language cannot be determined from the infant's form were considered as one. For example, consistent with Deuchar and Quay's (2000) report, M produces the form [ba] at 1;3 to represent *button*, but it is unclear whether she has acquired the English token of the word or the similar Spanish token *botón*. The authors mark this single form as representing both the English and the Spanish form, and then go on to report two further acquisitions, one of *button* and one of *botón*, thus accounting for three separate words in her early lexicon. However, in the present analysis only the first form was included in the dataset, as all further language-specific differentiation between the two forms (first produced as [bʌʔ] and [bɒn] respectively) was considered to reflect wider phonological and lexical changes taking place in the infant's output which are not relevant to this analysis. Vocabulary was considered in 10-word 'bins', providing a snapshot of each infant's lexicon over developmental as opposed to chronological time. Acquisition of both OW and CW forms was considered within the framework of the infants' wider lexical and phonological development.

All OWs are considered in this analysis, although it must be acknowledged that the definition of an onomatopoeic form may differ across the datasets. Indeed, some more idiosyncratic forms may not have been recorded in the parents' accounts, depending on their interpretation of what constitutes a 'word'. Furthermore, in some cases we observe the production of idiosyncratic onomatopoeia which are individual to particular infants or languages (for example, *houpy* to represent a swinging action in P's data), which may have been created by the infant based on the sound the object in question makes – these are termed 'innovative OWs' here.

2.4 OWs in the early output

2.4.1 Overview

Wide discrepancies can be found in the nine infants' approaches to OW production: Laura produces only three OWs, while OWs make up 22% of P's output over the first-100-word period. For five of the infants, OWs constitute a high proportion (more than one eighth) of the first 100 words. These are considered here to be 'high producers' (Annalena, Keren, Hildegard, Maarja and P), while Kaia, Laura, M and Trevor's outputs include no more than 11 OWs (less than one eighth of the 100-word samples). They are thus considered to be 'low producers'.

One common feature of OW use found in the group of high producers is the distribution of OW acquisition over the first 100 words. The largest number of OWs is acquired within the first 20 words: by the 30-word point, OW acquisition drops from an average of over three to under two new OWs per bin. This can be seen in Figure 2.1, where mean OW acquisition of the high producers is compared with that of the low producers.

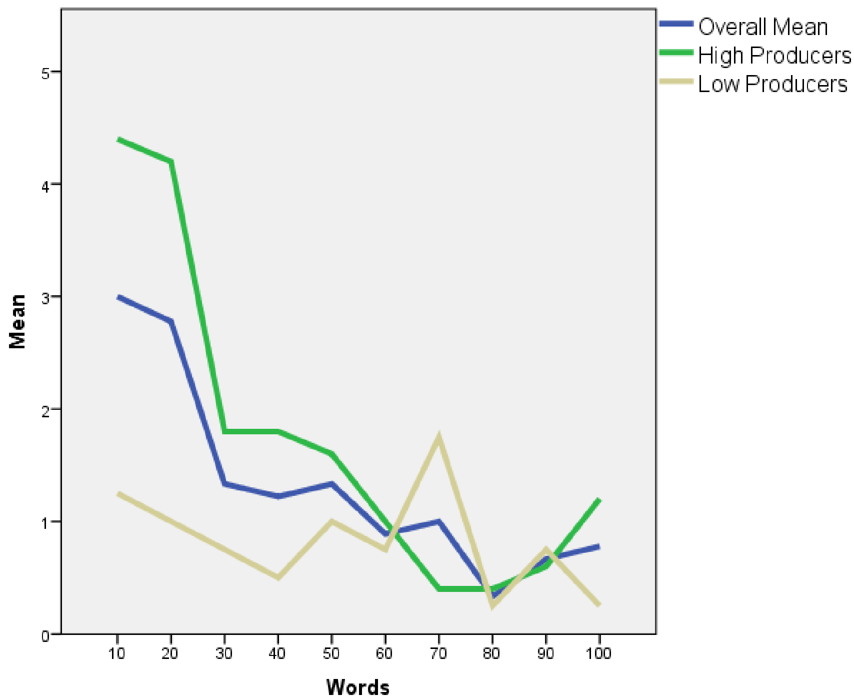

Figure 2.1: Acquisition of OWs in the first 100 words.

As Figure 2.1 shows, there is a consistency in OW acquisition for the low producers in this analysis, who are found to accumulate OWs steadily, at around one per 10 word bin, over the course of their first 100 words. Laura's OW production is very low throughout her dataset, and the three OWs that are produced within this period are spread evenly throughout, as she acquires no more than two OWs per 50 words (see Figure 2.2). Similarly, M has 11 OWs in her dataset, of which no more than 2 OWs appear in any 10-word bin. Trevor's OW acquisition peaks in the 20-word bin alongside the other infants in this study, but unlike the other infants, OWs account for only 9% of his first 100 words. Kaia, on the other hand, produces the majority of her OWs after the 50-word point, acquiring six of her nine OWs between the 50 and 70-word bins.

Of course, as the number of OWs produced in each 10-word bin decreases over time, so the number of 'regular words' (RWs: words which are neither OWs nor CWs) acquired in each bin increases. By the time that the infants have over 50 words in their lexicons, no infant is acquiring more than three OWs in any 10-word bin, and thus RWs are acquired at a higher rate. Indeed, for all infants except Kaia, the largest portion of OWs is acquired amongst the first 40 words, when these forms account for over a quarter of full word inventories for the five high producers. OW production remains high up until the 50-word point, when OWs account for between 22% and 38% of these five infants' lexicons. As Figure 2.2 shows, OWs constitute a large portion of many of the infants' full word inventories throughout the early word-learning period, and, as already discussed, even by the 100-word point they account for more than an eighth (12.5%) of all the high producers' lexicons – indeed, OWs make up more than 20% of the output for both P and Annalena over the 100-word period.

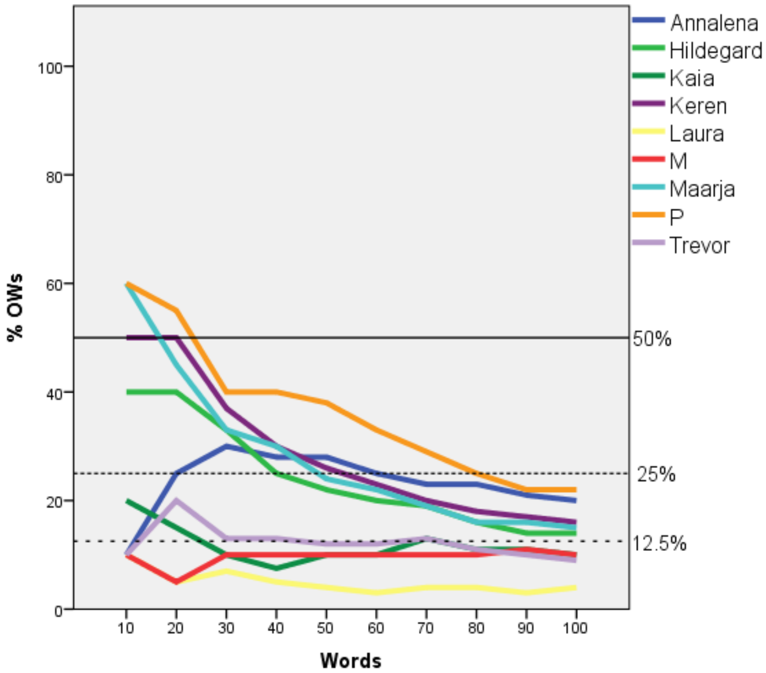

Figure 2.2: Proportion of OWs in the infants' first 100-words.

Figure 2.2 further highlights the contrast between the OW acquisition of the high and low producers. We can clearly see the low proportion of OWs in the four low-producing infants' outputs, with a general consistency in their OW acquisition over time as these forms make up an average of only 8% of their lexicons over the course of the analysis.

2.4.2 OW-CW combinations

The OWs in this analysis can fit into two distinct word categories: those which stand alone as abstract onomatopoeic forms, and those which have a corresponding noun functioning as a formal part of the language in question. *Bump* would be an example of the former, as there is no noun or alternative word (other than similar OWs such as *bang* or *crash*) to describe that which *bump* depicts. Animal and engine sounds such as *meow*, *vroom* and *moo* fall into the latter category, as these relate directly to the words *cat*, *car* and *cow*, and thus have conventional equivalents (CWs) which infants must acquire alongside the OW. When the nature of this latter category of OWs is examined more closely it seems that these words may have a functional role in language development.

Figure 2.3 shows the acquisition of 'functional' OWs in relation to that of their CW counterparts, taken from the data of the five high producers (the low producers' data will be considered below). Setting aside the nature of the individual word forms here, it is clear that every OW acquired by each of these five infants in the first 100-word period was first produced *before* the CW counterpart. That is to say that *meow* was always acquired before *cat*, *woof* before *dog*, *choo choo* before *train*, and so on. In all but five cases the CW equivalent was not acquired during the 100-word period, and in 11 cases the CW form wasn't acquired during the entire period of data collection, and cannot be accounted for at all in this analysis (marked as 'later acquisition' in Figure 2.3).

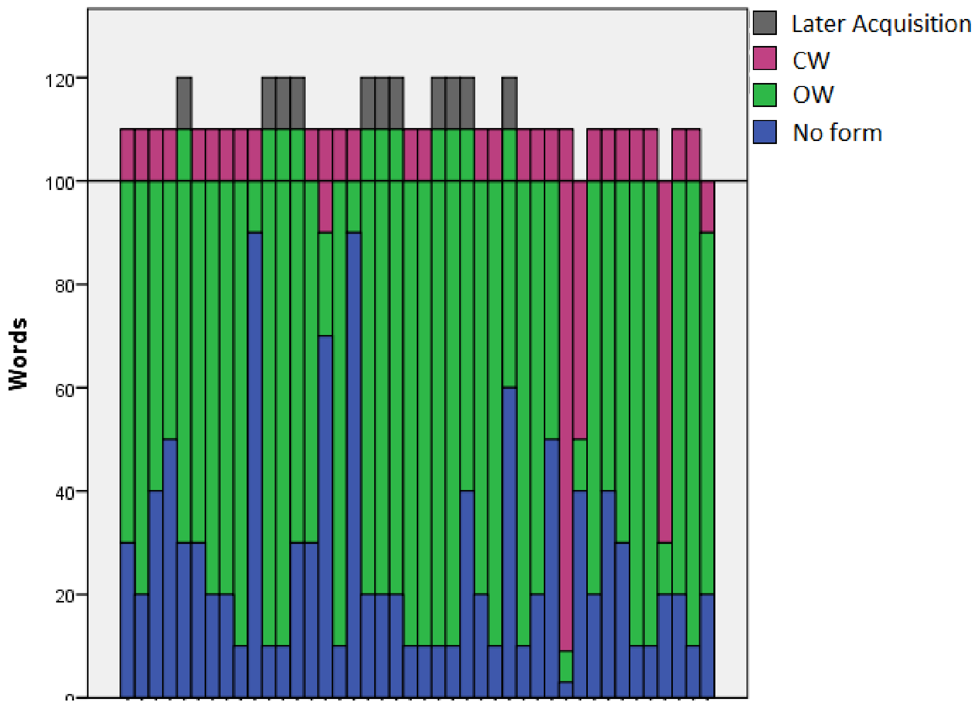

Figure 2.3: Transition from OW to CW in the high producers' datasets. CWs acquired after the marker at 100 words were acquired after the 100-word point – 'later acquisition' reflects words which are not accounted for in the full original datasets. Each bar on the Figure represents one word form in the dataset.

In the majority of cases Figure 2.3 shows a substantial delay between the infants' acquisition of the OW and their eventual first production of the CW equivalent: only seven CWs in the 42 OW-CW pairs are acquired within 50 words of the OW form, and almost two thirds (61%) are acquired outside the 100-word period. Here a marked contrast can be seen between OW acquisition, which is typical of the earliest stages of lexical development, and the acquisition of the CW, which occurs once the vocabulary is more established. Of course, these infants are already able to produce many RWs, and indeed they produce more of these forms than the OWs themselves at almost every point in the analysis. However, here we can acknowledge a potential role for OWs in early language development, as these forms appear to have more than just a subsidiary function in early infant interactions. Indeed, when an infant produces the form *woof woof* in the early stages of word production he may be referring to the dog itself, and not only to the sound that the dog makes. This proposal is confirmed for some of the infants considered here: Elsen (1994) remarks that Annalena produces *wauwau* 'woof woof' "for expressing withheld thoughts about dogs and other related objects" (p.311); many of Annalena's OWs are reported to be used referentially. Similarly, Maarja (Vihman & Vihman, 2011) is reported to use *aua* 'woof' in reference to 'doggie', while Leopold (1939) reports Hildegard's use of [m:] *yum* to "refer to food and eating" (p.97) at 1;0, and later its use as a verb, as in [wɑʊ ʔm] *Frau yum* 'lady yum' ('the lady is eating') as late as 1;8. The CW *eat* is not reported to appear in Hildegard's lexicon until 1;10, while [m:] is used actively to refer to 'something tasting good' as late as 2;0. We can conclude that when an infant has already acquired an OW they have no motivation to learn the corresponding CW, thus accounting for the long delay between OW and CW acquisition in this data. However, we must also ask why the infants produce the OW over the CW form across the board, if they both serve the same function referentially.

Indeed, the infants appear to prioritise the production of OWs over their CW equivalents, and furthermore, they can be found to 'create' their own OWs when an established form does not exist. Annalena produces [bɔa] in a hoarse voice as an onomatopoeic representation of *Krähe* 'crow', as well as a 'snuffling' sound for *Hase* 'hare' and a 'grunting' sound for *Schwein* 'pig'. There is no doubt that many of these forms – both lexicalised OWs and those that are innovated and thus idiosyncratic – are learned expressions, repeated from the caregiver's input. This has already been noted in Hildegard's use of [ʃ] to refer to cars and trains, and Elsen (1991) attributes Annalena's use of *wauwau* 'woof woof' to her grandmother, who used the OW form with Annalena while Elsen reports using only *Hund* 'dog'. However, Elsen (1991) does note that Annalena's production of the form [baʔbaʔ] for *quak* 'quack' is a "spontaneous onomatopoeic creation" (p.183, own translation), first used without any influence from adults to refer to both ducks and geese in the park at 0;11. Annalena is reported to produce this form intermittently in this context until a form more like the adult *quak* 'quack' appears – [gagak] – at 1;2.

When we compare these findings with OW production from the low producers, we can see some interesting contrasts. Firstly, only one of Laura's OWs relates to an animal sound – *bow-wow*; the other forms are onomatopoeia with no CW equivalent, such as *bump*. Furthermore, *bow-wow* is acquired after its conventional counterpart *dog*. Similarly, Kaia produces only one OW with a CW equivalent – *vroom* – but this appears as her third word, with *auto* 'car' following in the second 10-word bin. Kaia also produces three CWs (*kiisu/kitty, duckie* and *doggie*) in her 100-word dataset without acquiring the equivalent OWs in the same period. Trevor produces both *siren* and *vacuum cleaner* as [a:a:a:], and *razor* as [n:n:] or [m:m:], and appears to be relying on the creation of onomatopoeic 'protowords' in order to advance his word production. As can be expected, no CW equivalent is produced for these four OWs, but *kitty* is acquired before *meow*. M's OW production differs in approach yet again, as she acquires CWs *car* and *cat* in the first 10-word bin, and *train* a little later in the fourth 10-word bin, and in all three cases produces the equivalent OW much later on in the dataset. However, after this point her OW production appears to become more functional, as she acquires OWs *quack*, *bow-wow*, *oink* and *baa* (as well as *moo*, her first OW, produced in the first 10-word bin) all before the CW counterparts: *cow*, *pig* and *sheep* are not acquired until much later, after the 250-word point.

The low-producing infants provide the only OWs in the entire dataset that do not correspond to the pattern of acquisition shown in Figure 2.3. Again we see that these infants have a different approach to the use of OWs in their early output, which suggests that the OW production of the other five infants may somehow be functional: the contrast in approach between these two groups only underlines the potential bootstrapping effect that OWs may have in some infants' early outputs. In order to account for a role for OWs in language learning it is necessary to analyse these forms in more detail within the wider framework of the infants' phonological development.

2.5 Templates and OW production

Systematic phonological structures can be found throughout the infants' word production, as observed in accounts of templates in phonological development (Macken, 1979; Menn, 1971; Vihman, 2014; Vihman & Croft, 2007; Vihman & Velleman, 1989). These early prosodic patterns constitute simple structures such as CVCV with consonant harmony or reduplication, as in [bɛɪbɪ] *baby* (Laura) and [ta:ta] *daddy* (P), which are 'selected' systematically in line with the infants' output capacities. We also see some examples of forms which do not naturally fit the selected pattern and are thus restructured, or 'adapted', to do so. This is seen in P's 100-word sample, where a tendency to replace laterals with the liquid /j/ in his early word forms (as in *malý* 'small' [maja:] and *balon* 'ball' [baji] at 1;2) is later overextended to replace many other segments, adapting forms to fit a CVjV(C) template, as in *státi* 'to stand' [toji:] (1;3) and *zpivat* 'to sing' [pi:jat] (1;5).

It is well-established that infants' use of templates varies (Vihman & Keren-Portnoy, 2013), and indeed we see variability across the infants here. However, the limited period of development observed in this analysis reveals some noteworthy similarities across the nine infants' data in the establishment of consistent prosodic structures. The infants' individual patterns as well as consistencies across the group will be explored with OW production in mind, in order to identify whether OWs have a facilitative role in early lexical learning, as proposed above.

2.5.1 Methodology

In line with Vihman's (2015) approach to the prosodic analysis of bilingual infants' early word production, data from the nine infants in this study will be analysed to identify prosodic structures in their first 100 words. Any prosodic structure which occurs 10 or more times in the 100-word samples (including different variants of a single word type) will be considered here; the use of these structures will be assumed to reflect the individual infant's preferred output patterns, which correspond to their phonological and segmental capacity during the course of their first 100 words. Some of these prosodic structures may constitute the first use of templates in the infants' outputs: templates, as detailed in Vihman's analysis, evolve from the infant's preferred prosodic structures, whereby the 'pattern force' of their well-rehearsed routines brings about the over-selection of new words to fit that pattern so that we see the "overuse…of certain patterns in comparison with other children learning the same language" (2015, p.3). We can often observe segmental specification within these templates, as in P's palatal template described above, such that the infant appears to prefer particular phonemes and systematically replaces them in certain segmental positions. Template use is indicated by the use of adaptation to fit otherwise unmatched forms to the infants' preferred output patterns.

As in Vihman's analysis, differences in voicing, vowel quality and either manner or place of articulation (but not both) between word tokens will be considered insufficient to indicate differences in prosodic structure, while syllable structure, syllable onset, clusters and geminates will all be taken as indicators of prosodic variability. In line with Vihman's approach, this leads to differences in the number of word shapes included for each infant, as some of the datasets have a high degree of variability reported for each word form (e.g. Annalena, P), while others are limited to one or two word tokens to represent each of the infant's first 100 words (e.g. Keren, M). Furthermore, not all word tokens will correspond to a specific prosodic structure, and so the number of forms considered from each infants' 100-word samples varies from under 50 to more than 120 (see Figure 2.4).

2.5.2 Overview

When the infants' use of prosodic structures is analysed, clear consistencies can be seen across the data. Figure 2.4 shows the distribution of different prosodic structures across the infants' first 100 words, where we observe that consonant harmony (CH) plays a role in all nine infants' early outputs. Reduplication, monosyllabic C(V) structures and a minimal vocalic structure with a vowel onset ('vocalic structure', including V, VV, V: and V(V)C(C), as well as VCV) are also common: reduplication occurs in eight infants' early words, and C(V) and vocalic structures in seven and five of the nine infants' word forms, respectively. Finally, idiosyncratic patterns – labelled 'Own' in Figure 2.4, including Annalena's /l/-final template (CVl(V)), Kaia's long vowel (CV:) and geminate templates (CVC:V and VC:V), Keren's glottal structure (ʔV(C)), Maarja's front-rising dipthong pattern (CVi) and P's use of palatals, as described above – are specific to individual infants, demonstrating in each case the language- and infant-specific nature of even the earliest phonological patterns.

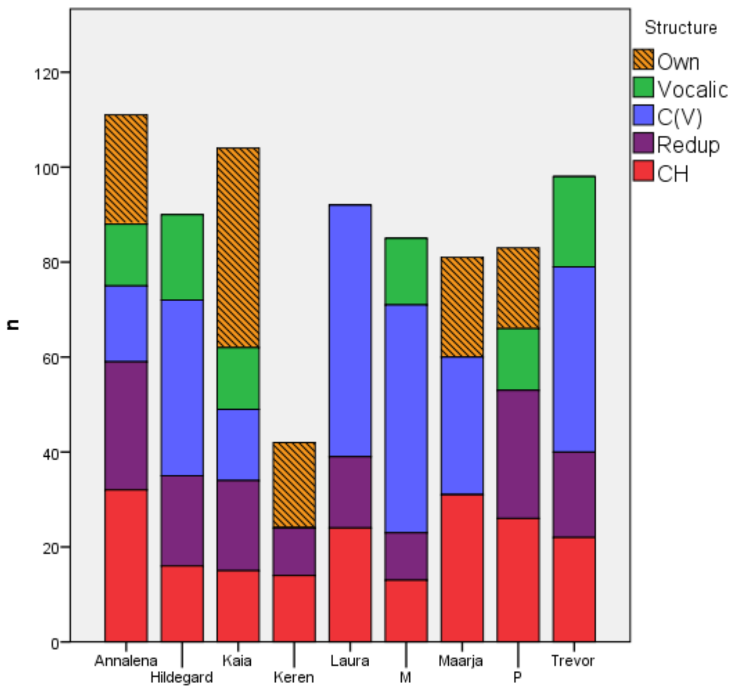

Figure 2.4: Distribution of prosodic structures across each infant's first 100 words.

For eight of the infants (all but Keren), more than half of the first 100 words fit within the same combination of four prosodic structures: consonant harmony, reduplication, vocalic and C(V). Again leaving Keren's data out of the analysis, we find that between 85% and 98% of the remaining infants' 100-word samples fit within the five prosodic shapes listed in Figure 2.4. Indeed, Keren is the only infant here whose word production is not heavily influenced by a specific prosodic pattern: between one quarter and a half of all other infants' first 100 words fit systematically into one specific prosodic structure, and in some cases, multiple structures. Invariably this includes consonant harmony (Annalena, Maarja, P), reduplication (Annalena, P) and C(V) (Laura, Maarja, Trevor, Hildegard and M), as well as the infants' individual templates (Kaia).

Figure 2.4 highlights some striking similarities between the output patterns of the infants in this study, despite the range of different languages represented in the data. This corresponds to findings from Boysson-Bardies and Vihman (1991), who observed similarities across the early prosodic structures of infants learning four different languages, despite identifying language-specific differences in infants' first words and in their babbling patterns. This is explained as being due to articulatory constraints, as infants implement a set of basic motoric patterns in the production of the more complex articulations in word forms. Phonological templates are suggested here with the proposition that infants produce early word forms based on "a selection favouring simpler articulatory patterns for words belonging to [their] everyday environment" (Boysson-Bardies & Vihman, 1991, p.316).

A small number of production patterns which are specific to individual infants begin to appear after the first words have been acquired. For example, P's preference for palatals is first observed in *auto* 'car' [aco], in the first 10-word bin, following which point palatals (including the glide /j/ and both voiced and voiceless plosives /c/ and /ɟ/) can be found in 51 of his first 100 words. As already described above, the systematic substitution of /l/ with [j] is predominant, featuring in every 10-word bin until the 100-word point, and the striking pattern force of this segmental preference is already apparent by the 25-word point at 1;2, when the previously accurate OW *lalala* 'singing' [laːlaːlaː] regresses to fit the palatal template as [jajajaː]. Similarly, Annalena has an approximant template whereby words featuring word-final or medial /l/ are selected for production in her lexicon; this first appears in the third 10-word bin with *Ball* 'ball' [bal̥ː] and continues to be used throughout her first 100 words as she adapts forms to fit this preferred output pattern. In the eighth 10-word bin this template accounts for six forms: *hoppola* 'oops' [bala], *Puppe* 'doll' [balˑa], *hallo* 'hello' (in reference to 'telephone') [halːo], *Lätzchen* 'bib' [lː], *Deckel* 'floor' [agˑl̩ːagˑl̩] and *Lappen* 'flannel' [balˑa], and we can see this pattern taking hold of her output as it becomes a template. Annalena's wider template-use will be discussed further in Chapter 3.

2.5.3 Output patterns and OW production

A consideration of the infants' OW production alongside their prosodic structures can help us to explain the wide use of these forms in early language. Tables 2.2-2.6 show the infants' use of OWs in terms of the five prosodic structures identified in this analysis. As is apparent from this data, most of the infants' OWs naturally fit the structures that are implemented widely across the dataset – that is that they are 'selected' to match the infants' preferred output patterns (Vihman & Croft, 2007). Bold highlighting demarcates those OWs which do not already match the prosodic shape of the target form, and thus have been adapted by the infant to match their preferred output patterns.

Reduplication

Reduplication constitutes a preferred output pattern for all of the infants except Maarja in this analysis, and is defined here as the repetition of full syllables across a word form. Both partial reduplication – whereby the vowel quality, articulation or voicing of one of the reduplicated segments is changed across the word form – and full reduplication are considered, and words are recognised as matching this pattern if they contain reduplication in the adult form as well as in the infant's output form. Reduplication has been seen as merely "a special case of consonant…harmony" (Smith, 1973, p.165), but here these two phenomena will be considered separately, as reduplication constitutes a particularly relevant feature of OWs, which differs from that of consonant harmony.

In Table 2.2 we observe minimal phonological discrepancies between target and output form, such as consonant or vowel assimilation (as in Laura's production of *bow-wow* as [waʊwaʊ] and Annalena's *bimbam* as [bambam]), cluster reduction (Hildegard's *klingelingeling* to [lilili]) and the mismatching or omission of consonants from the infant's form (as in Annalena's production of *quak* as [baʔbaʔ], which in fact is reported as an innovative OW and not as an attempt at the conventional OW *quakquak*). However, in all cases the infant forms are largely accurate, and all stay true to reduplication, which is in free variation with non-reduplication for most of these OWs (i.e. *quack* and *quack quack* are both acceptable variants of OW DUCK). As shown in Table 2.2, all cases of reduplication in the infants' OW production match this prosodic structure in the target form.

Table 2.2: Reduplication structure in the infants' OWs.

Infant	OW forms		
Annalena	*piep piep*	'squeak'	[pɪˑpɪˑpɪˑ]
	vafvaf	'woof'	[vava] [wawa]
	piepiep	'tweet'	[pɪˑpɪˑpɪˑ]
	bimbam	'bang'	[bambam]
	bingbong	'dingdong'	[mɔmmɔm]
	kikeriki	'cock-a-doodle-doo'	[kɪːkɪːç]
	quakquak	'quack'	[baʔbaʔ]
	mjammjamm	'yummy'	[namnam]
	miau	'meow'	[mimi]
	killekille	'tickle'	[kɪɫˑəkɪɫˑə]
Hildegard	*piep piep*	'squeak'	[pipi]
	mjammjamm	'yumyum'	[mjamjam]
	bimbam	'bang'	[biba]
	woof		[wuwuwu]
	natt natt	'quack quack'	[natnat]
	klingelingeling	'dingaling'	[lilili]
	choo-choo		[dudu]
	ticktack	'tick tock'	[tita]
Kaia	*choo choo*		[tuːtuː]
	tipa tapa	'tap tap tap'	[taptaptap]
	kukk kukk	'cock-a-doodle-doo'	[kukukuku]
	kiigakaga	'swing-swing'	[kigakaga]
Keren	*tutu*	'choochoo'	[tutu]
	lala	singing	[lala]
Laura	*bow-wow*		[waʊwaʊ]
M	*bow-wow*		[bəʊwəʊ]
	brm	aeroplane sound	[bmbm]
	oink oink		[ɔɪŋɔɪŋ]
	choo choo		[tʃtʃ]

P	*kaka*	'quackquack'	[ka:kaka]
	kokokodak	'cluck cluck'	[ko:koda:]
	lalala	'lalala'	[la:la:la:]
	huhu	'whoo'	[hu:hu:]
	bu bu	'moo moo'	[bu:bu:bu:]
	pipi	'tweet'	[pipi]
	cililink	'tingaling'	[sililin]
	chrchr	'grumph-grumph'	[xuxu]
	tudu	'honk honk'	[tidi:]
	tiktak	'tick tock'	[tsitsa]
	kutululu	'rolling ball'	[tutululu:]
	tuktuk	'taptap'	[tuttu:t]

Reduplication is found in all but Trevor's OWs, and is identified in Maarja's OW production[2], despite this structure not being sufficiently represented in her first 100 words to count here as a preferred output pattern. In Trevor's OW production, reduplication is widely used among his first 100 words, accounting for 18 of his earliest word forms, but the lack of OWs in his output leaves little opportunity to use this structure in this context. Almost all targets in Table 2.2 contain full reduplication, often in free variation with an unreduplicated equivalent (i.e. *moo* versus *moo moo*). Here we see how the structure of OWs matches the infants' preferred output patterns and thus also their articulatory capacity – indeed, as Figure 2.4 shows, reduplication is widely used and therefore well established in most of the infants' early outputs.

Consonant harmony

Consonant harmony is the most common prosodic structure in the infants' early outputs, and is the only output pattern found in all nine datasets. Here we consider full consonant harmony, whereby a consonant is fully assimilated to another consonant in the word (as in Hildegard's *kritze* 'hair brushing' [titsə]), and partial consonant harmony, whereby the place or manner of a consonant is assimilated across the word form (such as Kaia's production of *kõll* 'clink' as [tɤn:]). As with the examples of reduplication, we can see here that many OWs already contain consonant harmony, and moreover we find consonant harmony in many forms which also contain reduplication, marked in

[2] Found in *aua* 'woof-woof' [awawa], *uhuu* 'owl sound' [ʔuʔu:], and *ding dong* [dɪŋdɑŋ].

Table 2.3 with an asterisk (*). As with reduplication, words are considered to match this pattern if they contain consonant harmony in both the target and the output form.

Similar to the analysis of reduplication above, we find that many of the infants' forms are selected for consonant harmony in the target word (Table 2.3). In the case of *bong* (Annalena), *kritze* (Hildegard) and *kutululu* (P), we find the addition of consonant harmony in the child forms when it is not present in the target, but here we must consider the infants' phonological abilities at this stage, and the fact that [t] is a valid replacement for /k/ and [m] for /ŋ/ when these segments are limited in the output: neither Annalena nor Hildegard has yet produced any velar consonants and P only produces [k] in words with no differing consonants (*kuk* 'peep-o' [kuk], *kaka* 'quack quack' [kaka]). These forms are thus also considered to be selected for this structure. In many cases the presence of consonant harmony is automatic in a word form due to the use of reduplication, which we find in seven of the 21 target forms. Those forms which contain full reduplication of two syllables are not counted again in this analysis of consonant harmony, but note that the forms that appear in both Tables 2.2 and 2.3 contain differing consonants in the target (i.e. are partially reduplicated), and thus present segmental challenges that can be addressed through consonant harmony: these include voicing changes (*tudu* 'honk honk' – P) and changes in place or manner of articulation across the word (*tuktuk* 'taptap', *cililink* 'tingaling' - P, *bimbam* 'bang' – Annalena). While the reduplication structure alone can account for the infants' production of these OWs, the presence of consonant harmony in the target form should also be noted as an important feature.

Table 2.3: Consonant harmony structure in the infants' OWs.

Infant	OW forms		
Annalena	*brum*		[bvːm] [ʙm]
	*bimbam**	'bang'	[mamˑam]
	bong	'dong'	[mɔm]
Hildegard	*kritze*	hair brushing	[titsə]
Kaia	*nämmi/yum*		[mæmː]
	boom		[bomː]
	kõll	'clink'	[tɤnː]
	*kiigakaga**	'swing-swing'	[kigakaga]
Maarja	*mõmmi*	'teddy' (babytalk)	[mamːi]
	boom		[buːm]
	tudu	'sleep' (babytalk)	[tudu]
M	*quack/cuac*		[kak]
P	*tudu**	'honk honk'	[tidiː]
	*tiktak**	'tick tock'	[tsita], [tsitsa]
	bumbac	'bang bang'	[bumbaː]
	*cililink**	'tingaling'	[sililin]
	*tuktuk**	'taptap'	[tuttuːt]
	dutsduts	'bump'	[duts]
	*kutululu**	'rolling ball'	[tutululuː]
Trevor	*meow*		[maʊm]
	boom		[maʊm]

From this data it could be posited that the pre-specification of consonant harmony and reduplication in OWs makes them more learnable for young infants, as these two prosodic patterns are frequently found in infants' early words: all OWs in Tables 2.2 and 2.3 contain consonant harmony or reduplication (or both) in the full form, enabling infants to produce them with a high level of accuracy (i.e. 'selecting' these OWs) even in the very earliest stages of word production. Furthermore, both of these phenomena are reported as being particularly prominent in early language production (Smith, 1973; Vihman, 1978): Ingram (1974) discusses both reduplication and consonant harmony (referred to in Ingram's paper as 'assimilation') as common phonological features of early child language, and Smith (2004) goes so far as to refer to consonant harmony (and thus, by his own definition, reduplication) as a 'universal' feature of child language (p.303).

C(V) structure

When the infants' other preferred prosodic patterns are considered, much more variability can be found both between and within infants. The C(V) structure is the only other output pattern to be used by all infants in the production of OWs, incorporating simple phonological shapes featuring a single consonant combined with a single vowel or dipthong, including CV, CVV, C: and CV: forms. Infants' production of CV syllables in early language development has been widely discussed in the literature, and has been posited as another 'universal' of language development (Kent, 1992) owing to infants' early physiological predisposition towards mandibular movements, which bring about canonical babbling (Kent, 1992; Vihman, 1992). In Vihman's (1992) study of early syllable production across four languages, it was found that more than half of the infants' first word types constituted CV syllables that had been practiced in the pre-linguistic phase, highlighting the continuity between pre-linguistic vocalisations and the development of the early lexicon. The data observed from the infants in this study provides further evidence towards the ubiquity of the CV structure in early production.

Many OWs are reported here as naturally fitting the C(V) structure (Table 2.4), but we can see that some of these forms do not match the target as accurately as those observed in Tables 2.2 and 2.3 (e.g. Annalena's *tööt* 'toot' [bə] and Laura's *bump* [bu]). However, in all but two cases (*cock-a-doodle-doo,* Maarja, and *hush*, Trevor) this mismatch involves the omission of the final consonant, which is phonologically restricted (and in many cases omitted completely) across all nine infants' early words. This suggests that these forms have not been adapted according to Vihman's (2015) definition of this term, but are instead selected in line with the infants' early output capacity: as word-final consonants are not yet a phonological feature of many of these infants' outputs, the omission of this segment does not reflect the adaptation of these word forms.

Table 2.4: C(V) structure in the infants' OWs.

Infant	OW forms		
Annalena	*meh*	'baa'	[mɪː] [meː]
	tööt	'toot'	[bə] [bɪ]
	puh	'phew'	[bː]
		cawing sound	[bɔa]
	muh	'moo'	[mː]
Hildegard	*moo*		[muː]
	pieks	'prick, sharp'	[by]
	mmm	'food'	[mː]
	sshh		[s]
Kaia	*vroom*		[vːvːv]
	moo		[muː]
	pai	'pat', 'nice'	[pa]
Keren	*mu*	'moo'	[mu]
	dio	'giddy up'	[dio]
	myau	'meow'	[miau]
	eeyore	donkey sound	[ʔia]
Laura	*bump*		[bu]
Maarja	*moo*		[muː]
	baa		[ba]
	cock-a-doodle-do		**[ka]**
	whee		[wiː]
	ka	goose sound	[ka]
M	*moo*		[mː]
	quack/cuac		[ka]
	miau	'meow'	[maʊ]
	beh	'baa'	[be]
	yum/nam		[mː]
P	*bebe*[3]	'baa'	[beː]
	muk	'shh'	[mu]
	bac	'bang'	[baː]
	mnau	'meow'	[naː]
Trevor	*razor*	razor sound	[nːnː] [m;m;]
	boom		[baʊ]
	hush		[ʃː]

[3] As in English *baa/baabaa*, reduplication in this form is assumed to be in free variation.

Again we see that the majority of forms in Table 2.4 (97%) do match the infants' preferred prosodic shapes. Indeed, in most cases the omission of the word-final consonant is the only feature of these forms which does not correspond to the adult target: in terms of syllabification and even the onset consonant, the majority of these forms remain faithful to the target form.

It should also be noted that many of the words listed in Table 2.4 are the same output forms produced across all of the infants: this structure accounts for five infants' production of *moo*, four infants' productions of each of *baa* and *meow* and three of *shh*, as well as two of *yum/food*. We also see *boom/bump/bang* appearing in three instances, with the final consonant omitted in all cases in order to fit the prosodic shape CV. The phonological similarities observed in OWs across languages appears to facilitate their learnability in the early stages of many infants' lexical development.

Vocalic structure

The final prosodic pattern to occur consistently across the infants' outputs is the vocalic structure, which accounts for prosodic shapes that are dominated by vowels, either through the absence or omission of any consonants in the form, or by a vocalic or dipthongal onset. This includes the structures V(V), V:, (V)VC, VCC and VCV. Words were considered to match this prosodic pattern if the target form matched one of these five prosodic shapes, or if the target onset consonant was a glottal stop or fricative and thus less salient auditorily. As shown in Table 2.5, most OWs in this category match the target form appropriately

Table 2.5: Vocalic structure in the infants' OWs.

Infant	**OW forms**		
Annalena	*aua*	'ouch'	[awa]
P	*houpy*	swinging sound	[o:pi]
Trevor	*ow*		[aʊ]
	woof		**[aʔ]**
		siren sound	**[a:a:a:]**
		vacuum cleaner sound	**[a:a:a:]**

This structure is much less productive, accounting for only five OWs across three infants' data. From Table 2.5 we can see that OWs do not generally match this pattern automatically, and aside from *ow/aua* 'ouch' we see no lexical trends in its implementation. Trevor makes the most use of this structure in his first 100 words, as we see the imitation of sounds from his surroundings using one of his preferred output patterns. In the case of his *siren* and *vacuum cleaner* sounds we can assume that these OWs are innovative, and we can thus speculate as to the role of intonation in Trevor's use of this pattern: it seems likely that the long vowels used to reflect specific meaning may have been differentiated – and thus made recognisable to the diarist – by the appropriate intonation. However, without further evidence it is not possible to resolve this here.

Individual structures

The remaining prosodic structures that occur as preferred output patterns in the dataset are not consistent across infants, but are instead specific to the individual infant and the features of the language that they are acquiring. Keren produces many glottal stops and glottal fricatives, as is typical of Hebrew, for example, and we see a preference for palatals in P's word forms, similar to the palatalization features observed in the early word forms of Grzenio, acquiring Polish (Szreder, 2013). Pačesová (1968) comments on P's use of palatals, noting that /j/ is "very well learned" (p.48), even from the very outset of word production. Maarja's CVi structure, termed here the front-rising dipthong pattern (in line with Vihman, 2015), has also been referred to by Vihman and Vihman (2011) as a palatal template, and has been found to be proportionate to the use of palatals in Estonian IDS.

Phonological preferences can also be found which do not specifically relate to the ambient language. For example, Annalena demonstrates a preference for liquids, which can be attributed to a prominent use of /l/ during her babbling phase (Elsen, 1991). Bias towards liquids and glides in early word production has been reported in the literature, with a preference for /l/ reported in Laurent's early word forms (Vihman, 1993a) and palatal templates identified in English, as in Priestley's son Christopher (1977), and in French (Smith, 2011). Both Annalena's preference for /l/ and P's preference for palatals lead to the use of templates in their later word production (see Chapter 3 for an analysis of Annalena's templates), but even in the first 100 words we can see how these prosodic structures shape their early outputs. Table 2.6 demonstrates Annalena, Maarja and Keren's production of OWs which match their preferred individual structures. As we see no evidence of P's palatal template in his OW production this is not considered here.

In Annalena's data we can see another example of a form which fits the reduplication structure (marked with *). In Keren's case, the glottal-initial pattern could also correspond to the vocalic structure discussed above, but here we differentiate between these two possible output patterns, as Keren produces only five words with a vocalic structure, compared with 27 correctly-produced glottal-initial words. Here we consider her individual structure to be phonologically specified due to its abundance in Keren's early output, while the vocalic structure does not appear to play a role. Similarly, Maarja can be found to differentiate the CVi pattern and the C(V) pattern, as both are equally common in her first 100 words.

Table 2.6: Individual structures in OW production.

Infant	OW forms		
Annalena	*killekille**	'tickle'	[kɪɫˈəkɪɫˈə]
/l/ structure	*hoppola*	'whoops'	[bala]
Keren	*haw*	'woof'	[haw]
Glottal-initial structure	*ham*	'yum'	[ham]
		'eeyore'	[ʔia]
	hita	'walk'[4]	[hita]
	hupa	'bump'	[hupa]
		'disgust'	[ʔix]
Maarja	***pai***	**'nice'**	**[ʔai]**
CVi structure	***kõll***	**'clink'**	**[ʔɤɪ]**

Again we see a dominance of forms which match the adult target in Table 2.6: only two forms, both in Maarja's data, do not correspond to the target form, and then it is only *kõll* 'clink' which does not contain the front-rising dipthong that defines this pattern. It cannot be claimed that the accuracy of these forms is due to the general phonological features of OWs, as the three structures analysed here are specific to each infants'

[4] Defined as 'babytalk' in Dromi's (1987) account.

individual output and language experience. However, this only supports the evidence to suggest that infants' OW production, and indeed their early word production in general, is dependent on the constraints and preferences that we find throughout the individual infants' outputs.

2.5.4 Summary

An analysis of nine infants' preferred output patterns has revealed both trends and inconsistencies across the infants' data. More than three quarters of the infants' OWs (77%) correspond to only three prosodic structures, all of which have been identified as the most common preferred patterns in these infants' early outputs. All three patterns have also been identified as 'universals' of infant language production (Kent, 1992; Smith, 2004) owing to their presence in many infants' early outputs across a number of different languages.

It is unsurprising that we find reduplication, consonant harmony and the C(V) structure so often in infants' early words, since these three simple patterns are less of a burden in terms of memory, planning and articulation. The fact that OWs tend to fit these patterns without any real need for adaptation makes them not only learnable but also producible, allowing us to postulate one possible reason as to why we see so many of these forms in the earliest stages of lexical development.

The lack of correspondence between OW production and the vocalic and individual structures supports this point further, since both of these patterns were influential in the overall dataset, albeit less so than the three structures discussed above. While we still see the production of OWs in line with the infants' preferred output patterns, the large discrepancy between the productivity of these two structures and that of the patterns observed in Tables 2.2-2.4 reveals a potentially important consistency across the prosodic shapes found in OW forms. Indeed, patterns which are common in the infant's general output do not necessarily lead to their use in OW production: the specificity of OW forms and their prosodic shapes across languages appears to facilitate their learning in this regard, as their simple phonologies coincide with the so-called 'universal' features of early language production, corresponding to the rudimentary prosodic shapes that are found across many infants' early outputs.

2.6 Prosodic structures and CW production

As we have already seen in Figure 2.3, in all instances of OW acquisition the high-producing infants acquire their OWs before the equivalent CWs. Figure 2.4 shows us that the prosodic patterns identified in OW production (Tables 2.2-2.6) can also be found throughout the infants' first 100-word samples – implemented in the production of RWs as well as OWs – and so it can be assumed that they also occur in the acquisition of CW forms. Furthermore, the low-producing infants can often be found to acquire CW forms prior to acquiring their OW equivalents. The nature of these CWs should therefore be analysed in order to determine how the infants' preferred output patterns might be implemented here.

An overview of the infants' CW forms can be seen in Table 2.7, where it is clear that their preferred output patterns are found in CWs as well as OWs: of the 38 CWs listed, only one form fails to match at least one of the infants' prosodic structures (Hildegard's *duck*). Furthermore, contrary to the OW analysis, we see that the majority of the CWs do not match the infants' output patterns in the target form, and are thus adapted by the infant (highlighted in bold) to match their preferred prosodic shapes. Indeed, 66% of CW forms are adapted in at least one token, which contrasts sharply with the proportion of adapted OWs discussed above, where 86% of the infants' OW forms were accurately produced, or 'selected'.

Table 2.7: Template use across OW and CW forms in the infants' first 100 words.

Infant	OW				CW			
Hildegard					*kitty*		**[diti]**	**CH**
					auto[5]	'car'	[ata]	Vocalic
	natt-natt	'quack'	[natnat]	Redup.	*Duck*		[dak]	None
Kaia					*kiisu/kitty*		[kiːtɔ]	CH
	vroom		[vːvːv]	CV	*auto*	'car'	[atːo]	Geminate
					doggie		**[ʌkiː]**	**Vocalic**
					duckie		[dʌki] **[dʌti]**	None, **CH**
Keren					*auto*	'car'	[ʔoto]	Glottal
Laura					*car*		[ka] **[kaka]**	C(V), **Redup.**
	bow-wow		[waʊwaʊ]	Redup.	*doggie*		**[gɔgɪ]**	**CH**
					kitty		**[ki]** [kɪdɪ]	**C(V),** None
					horsie		**[sɪ]**	**C(V)**
					cat		**[ka]**	**C(V)**
					bird		**[bø]**	**C(V)**
					clock		[kɔk]	CH
Maarja					*duck(ie)*		**[da]**	**C(V)**
					horse		[os]	Vocalic
					bee		[biː]	C(V)
					kass(i)	'cat'	**[asːi]**	**Vocalic**
					notsu	'piggie'	**[nunu]**	**CH**[6]
					birdie		**[tiː]**	**C(V)**
					lind	'bird'	**[nɪn]**	**CH**
					clock		**[kɑ]**	**C(V)**
					car		[ka]	C(V)
					auto	'car'	[aʊto]	Vocalic

[5] Hildegard uses the form [ʃː] to refer to cars and other locomotives. This appears to be an overextension of *choochoo*, and is thus not considered here as an equivalent OW for *auto*.

[6] Maarja does not have a reduplication structure in her first 100 words, but consonant harmony can be found throughout.

M					*car, carro*		[ka]	C(V)
					clock		**[ka]**	**C(V)**
	meow		[maʊ]	C(V)	*cat, gato*		**[ka]**	**C(V)**
	quack, *cuac*		**[ka]**, [kak]	**C(V)**, CH	*duck*		**[da]**	**C(V)**
	bow-wow *woof*		[bəʊwəʊ] [wʊf]	Redup. None	*dog*		**[da]**	**C(V)**
	choo choo		[tʃtʃ]	Redup.	*train, tren*		**[tu]**	**C(V)**
	moo		[m:]	C(V)	*vaca*	'cow'	**[aʔa]** [vaʔa]	**Vocalic**, None
P	*tudu*	'honk'	[tidi:]	CH	*auto*	'car'	[aʊto]	Vocalic
	mnau	'meow'	[na:]	C(V)	*čiči*	'kitty'	[tsitsi]	Redup.
	bebe	'baa'	[be:]	C(V)	*beránek*	'lamb'	**[bejani]**	**palatal**
Trevor					*duck*		**[dʌ]**	**C(V)**
					clock		**[kæ]** [kæk]	**C(V)**, CH
	meow		[maʊm]	CH	*kitty, cat*		**[kiki]** [kɪt]	**CH**, None
					dog		**[gʌ]** [dʌ]	**C(V)**
					car		[kæ] [ka:]	C(V)
					bird		**[akla]**	**Overext.**[7]

[7] [akla] is used for sky-related objects including *cloud* and *aeroplane* as well as *bird*; is it likely that we see an over-extension of *cloud* here. All three forms are acquired within the same 10-word bin.

When OWs and CWs are compared in this way, it is unsurprising that the trajectory of OW and CW acquisition follows such a distinct and consistent pattern across the high producers' lexicons. OWs are acquired earlier than their corresponding CWs in all cases because they fit most appropriately to the well-established prosodic structures in the infants' output. Indeed, in the majority of cases the CW forms do not match the target structures, and must be adapted to fit the infants' output patterns. As explained by Vihman (2015), word adaptation takes place once the infant is ready to attempt more challenging words, when he or she has gained experience of producing existing prosodic routines through selecting forms which already fit these shapes. We can see this trajectory in the early use of OWs across these infants' data, before they later come to adapt CW forms to fit their established output routines.

However, this does not explain the trajectory of OW/CW acquisition found in the low-producing infants, who show no consistent trend in this domain. When we look at these four infants' data in Table 2.7 we see a high use of CW adaption: 79% of CWs are adapted to fit their preferred output patterns, and so it is clearly not that these infants' CW forms provide a 'better fit' than the equivalent OWs. Furthermore, OW forms are in most cases selected to fit the appropriate prosodic structure: all nine of these infants' OWs listed in Table 2.7 are selected in at least one token. This suggests that the OWs provide a 'better fit' to the infants' preferred output patterns, and if we consider the OW and CW forms to be 'competing' as potential labels for specific referents, it seems that the OWs should be selected before their corresponding OWs. However, a close-up view of the phonological properties of these infants' early outputs provides some further insight into their acquisition of CW forms. Both Laura and Trevor are found to have a preference for /g/ and /k/ in their first 20 words, as well as prosodic preferences for the structures that are identified in their earliest CWs: Laura's *doggie* [gɔgɪ] is preceded by the acquisition of *car* [ka], [kaka], and followed by two more /k/-initial forms (*cookie* and *kitty*) before the 15-word point. Furthermore, 13 of her first 20 words contain full consonant harmony or reduplication (including *baby* [beɪbɪ], *mommy* [mɔmɪ], *banana* [baba], *daddy* [dada], as well as *bow-wow* [waʊwaʊ]). At the point that *bow-wow* is acquired, shortly after *doggie*, we see a new preference for bilabials, as this form is acquired in a cluster of six bilabial-initial words. Trevor shows an even stronger preference for velar plosives throughout his earliest words: 30% of his first 30 words are produced with a [k] or [g], while in his first 20 words all six forms containing a velar stop in the target are realised with a velar stop. However, the strongest pattern observed in Trevor's early output is the application of the CV structure, which accounts for 9 of his first 20 words and leads to the adaptation of CWs such as *duck* [dʌ] and *dog* [gʌ], [dʌ], [dæ]. These two infants' early outputs therefore appear to be driven by phonological as well as prosodic cues: Laura demonstrates an early preference for velar-initial words, as well as consonant harmony and reduplicated forms, which are applied in the production of CWs *doggie* and *kitty*, and, shortly afterwards, OW *bow-wow*. Similarly, Trevor selects forms which contain velars in the target, and adapts many forms to fit his preferred CV output pattern.

M's CW production differs from that of Trevor and Laura as she has many more OW-CW pairs in her first 100 words, among which some OWs are acquired before the corresponding CWs and some after. However, it is clear that M's early output is driven by a strong preference for a CV pattern, and this also appears to dictate her acquisition of OW and CW forms. Sixteen of M's first 20 words fit a strict CV structure, and in fact her first OW *moo* [m:] is one of the four forms which does not. Alongside this preference for CV forms, however, we can also see a clear phonological influence on M's early output, as the same CV syllables are used to refer to a range of appropriate target words: in the first 10-word bin we see the acquisition of *car/carro*, *clock*, *casa* 'house' and *cat/gato*, all produced as [ka], and in the second 10-word bin *papá* 'daddy', *pájaro* 'bird', *panda*, *button/botón*, *ball/bola*, *más* 'more', *baby/bebé* and *M* (child's name), all produced with a bilabial stop followed by an open vowel. M's OWs and CWs fit with these patterns, as her output is highly restricted and thus few words are considered to be selected in her early production. Prolific use of the CV template in her early words drives word adaptation on a broad level, whereby similar-sounding targets are acquired simultaneously, all represented by the same or similar output forms – the acquisition of *quack/cuac* [ka] alongside *clock*, *cat/gato* and *car* is evidence of this. In later production, a more varied consonant inventory and a wider set of prosodic structures enables the acquisition of OW forms, which, in contrast with the high producers' data, are phonologically more complex than her early-acquired CW forms: compare *dog* [da] with *bow-wow* [bəʊwəʊ] or *woof* [wʊf], *cat* [ka] with *meow* [maʊ] and *train* [tu] with *choo choo* [tʃtʃ] in relation to the early acquisition of *moo* [m:] in contrast with *vaca* 'cow' [aʔa] [vaʔa].

Kaia's data differs from the other low-producers yet again, as she has very few CW forms in her first 100 words. Her very first word, *kitty/kiisu*, is a CW, which is produced in a whisper three months before the second word appears in her lexicon, and is considered by Vihman (2015) to be a progressive idiom. The next CW in her lexicon, *auto* 'car', is acquired after the equivalent OW form, but an important detail to note here is that this form fits her geminate VC:V structure, and is acquired just as this is beginning to gain prominence in her output. *Auto* [at:o] presents only the fourth example of this structure in her 100-word sample, but 11 of the next 16 new words fit this pattern. We see strong systematicity throughout Kaia's 100 words, with a high proportion of the first 50 words fitting to a disyllabic VC(:)V shape, or CVC(:)V with consonant harmony. A prosodic perspective alone would lead us to assume that this may prompt the acquisition of reduplicated OWs such as *woof woof* and *quack quack*, but again, if we observe the phonological trends in Kaia's output at this time, we are able to account for the acquisition of the CW forms over the OWs in both of these cases. Apart from the progressive idiom *kitty*, *doggie* [ɔki:] presents the first instance of /k/ in Kaia's output, and six of the following 12 new words also contain this phoneme, one of which is *duckie* [dɔki]. As with the other three low-producing infants, it seems that Kaia's lexical acquisition is driven by both phonological and prosodic cues, as new words are produced in line with her developing articulatory abilities. *Duckie* and *doggie* present similar challenges in the output, and consequently they are acquired only three words apart, and alongside words such as *kukub* 'falling' [kuk:u], *cookie* [kɔk:i] and *katki* 'broken, torn' [kat:ɪ] [tak:i]. A similar trend can be seen with the CV(:) template, which becomes dominant in the third 10-word bin, and accounts for one third of all words at this point. As in M's case, we see the contiguous acquisition of similar words throughout Kaia's data; for example, OW *moo* [mu:] follows the acquisition of *kuu* 'moon' [ku:] and *moon* [mu:].

2.7 Discussion

This chapter sought to identify how OWs might fit within nine infants' developing phonological systems. Overall, these forms were found to match the common prosodic structures that are observed in many infants' early lexicons, and so it seems that OWs may be typical of early production owing to their simple articulatory shapes which suit infants' developing production abilities: when reduplication, consonant harmony and C(V)structures are considered together, correspondences can be found between the infants' realisation of OWs. The majority of OWs occurring in this dataset naturally fit one (or all) of these three prosodic patterns; many are produced with reduplication (such as *woof woof* and *quack quack*) or constitute a simple CV structure that can be easily reproduced in infant speech as a single syllable (for example *moo* and *baa*). Indeed, when we look at typical OW forms such as *woof*, *baa*, *moo* and *bang* we can see that a combination of the same set of OWs are produced by most of the infants in this analysis, fitting one, two or all three of the common prosodic patterns identified in this data. Furthermore, these prosodic consistencies can be observed despite the differences in target language across the nine infants, and the commonalities in OW forms across the languages being acquired in this study lead to similarities in the infants' early lexicons. Similarities in the prosodic patterns produced across infants and the prosodic structures found across OWs may explain the presence of OWs in first-word data across a wide range of languages (Menn & Vihman, 2011, Appendix I).

Consonant harmony, reduplication and the CV syllable have been found to be especially prevalent in early infant word production (Ferguson, 1964; Ingram, 1974; Kent, 1992; Smith, 1973; Vihman, 1978; Vihman 1992), and the common presence of these features in OWs provides an appropriate fit in terms of the infants' preferred output patterns. Indeed, it could be proposed that these infants are selecting OWs over the conventional adult form, thus increasing word production while also working within the boundaries of their own output capacities. As *moo* has been found to be acquired before *cow*, and *woof* before *dog*, and so on, in all cases across the high-producing infants' datasets, it seems that the infants may be taking advantage of the multiple lexical options that are available to them where words with an OW equivalent are concerned. It is plausible that the infants would acquire an OW over its phonologically more complex CW equivalent, especially when the prominence of prosodic patterns such as reduplication, consonant harmony and C(V) syllables are observed in the output. Once an infant has a form to represent *dog*, *cow*, *sheep*, et cetera, there is no real motivation to acquire the CW form; in many cases the original datasets provide evidence that the infants are using OWs referentially in place of the corresponding CW, which explains the somewhat long delay that is often seen between the acquisition of the OW and the CW. This transition will be analysed more closely in the next chapter.

This brings us back to the wider pool of data, where we can now reconsider the low-producing infants' word production in light of these findings. These infants all make prominent use of the common prosodic structures observed here, and indeed reduplication, consonant harmony and the consonantal template are applied to some extent in all four infants' data: Laura's use of the C(V) structure is the highest across all infants (53%), while M and Trevor exhibit third (48%) and fourth-highest (39%) use of this prosodic pattern, respectively; in this sense, these infants' data fits appropriately with that of the wider study. However, a closer look at the phonetic features of the low producers' outputs demonstrates a preference for specific segmental properties within the established structures. In all four cases this was found to lead to the targeting of word acquisition on both a phonological and a prosodic level, thus prompting the early production of CW forms which fit both the preferred prosodic and segmental features of the infants' outputs.

Evidence from this analysis suggests that the simple phonological shapes of OWs often correspond to the infant's well-rehearsed output patterns, thus promoting the acquisition of OWs over CWs in many cases. However, this is defined by the individual infants' outputs and thus the same process can be found amongst those infants who scarcely produce OWs: Laura acquires *doggie* before *bow-wow* and Kaia and Trevor first produce *kitty* before *meow*, whereby the diminutive CW forms are 'competing' with a version of the OW which is perhaps less of a match to the infants' preferred output segments and structures. In these examples, both CW forms match the infants' phonetic preference for velar plosives, while adaptation is required to fit both OW and CW forms to the infants' preferred prosodic patterns. This contrasts with the other infants' approaches, who in many cases produce the reduplicated form of the OW *woof woof* (*vafvaf* [wawa] (Annalena), *woof woof* [wuwu] (Hildegard), *hawhaw* [hawhaw] (Keren), *aua* [wawa] (Maarja)) to fit these infants' reduplication structure. Finally, we can consolidate these findings through the consideration of M's data, whose prosodic patterns correspond most typically to the CW forms, and thus she can be found to acquire these words earlier than their equivalent OWs. However, in other cases the OW is better-suited to her output, and so this form is acquired before the CW. This evidence suggests a strong phonological motivation for the acquisition of OWs, which makes them no different from the rest of the vocabulary in the process of language acquisition.

The final aspect of this dataset that should be acknowledged is the wide variation in the infants' ages from acquiring their first to their hundredth word. Annalena is the youngest of the infants to begin word production, and the possible relevance of this to the high proportion of OWs in her output – the second highest across all of the infants – should not be overlooked. This is in direct contrast with Laura, the oldest of the nine infants, who produces the lowest number of OWs. Indeed, when the infants' ages are considered alongside their OW production, the youngest infants to begin producing words have the highest number of OWs in their datasets. Trevor and M are exceptions to this trend, but when age across infants is considered in comparison with OW use, a general tendency becomes clear (Figure 2.5). Indeed, even in a small sample of nine infants, and with Trevor and M's status as exceptions in this dataset, a Pearson product-moment correlation comparing the infants' age at onset of word production and number of OWs produced is found to be significant (r=-.675, n= 9, p=.046). Furthermore, this significance increases when Kaia's data is considered from the production of the second word, as her first word *kitty* appeared three months earlier and is reported by Vihman (2015) to be a progressive idiom (r=-.700, n= 9, p=.036). This is reflected in Figure 2.5.

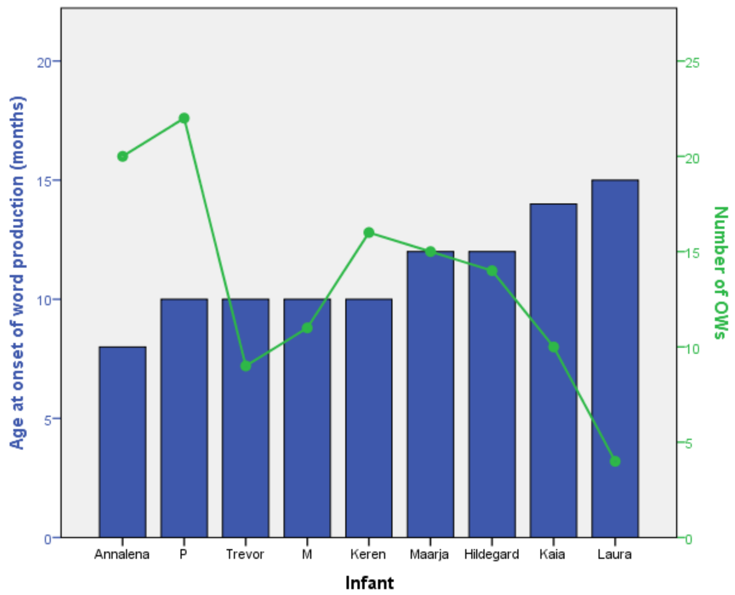

Figure 2.5: OW production in relation to age across the dataset.

Across seven of the nine infants observed in this dataset, the proportion of OWs in the first 100 words stands in relation to the age at which the infant begins word production. Cultural and behavioural factors are no doubt accountable for an infant's acquisition of OWs, and thus these differences may also relate to the infants' language experience, including interactions between the infant and their caregiver: Laura's increased language experience may present her with a broader lexical knowledge, allowing her to bypass the potential 'stage' of OW-production that we see amongst the younger infants. Younger infants may also have more exposure to 'baby-talk' words, including OWs, in their input, as mothers have been found to reduce the number of these words produced in the output as their infant's age increases, favouring instead more 'adult-like' word categories such as nouns and verbs (Kauschke & Klann-Delius, 2007). An analysis of caregivers' production of OWs in Chapter 6 will provide further insight into the role of the input in OW production, allowing us to observe the changing nature of the caregiver's OW-use over time.

2.8 Conclusion

This analysis has provided a broad overview as to the reality of OW acquisition in the early output. Trends observed across nine infants can account for a phonologically-motivated role for OWs in early word production, which is also dependent on various external factors such as the target language and the infant's age, as well as the systematicity of the infant's early input and the phonological preferences that each infant exhibits. It seems that OWs allow infants to select the form which is most suited to their limited production capacity, providing a second option in lexical acquisition which may be better-suited to the constraints of the early output. This allows the infant to expand their lexicon and gain production experience while also working within the confines of a minimal phonological inventory. These findings suggest a highly productive role for OWs in phonological development, as they allow for selection on a lexical level in order to facilitate production – an essential precursor to further language development.

Annalena's data provides a close-up view of the implementation of OWs in early word production, with a range of examples of onomatopoeic imitation whereby Annalena is reported to 'create' OWs through the imitation of sounds in her environment. We will now look more closely at Annalena's data specifically in order to consider her OW production in relation to the changes taking place throughout her phonological development.

3. Diary analysis II: A longitudinal analysis of Annalena

3.1 Introduction

Following the analysis of early prosodic patterns and OWs in Chapter 2, a longitudinal analysis of production data from one infant will provide a detailed bottom-up view of early OW production, tracing the acquisition and decline of these forms with reference to the wider phonological changes taking place in the infant's output. Elsen's (1991) diary account of her daughter Annalena provides ample evidence for the use of onomatopoeia in phonological development. Indeed, we have already acknowledged the large number of OWs produced in Annalena's early lexicon, and so her data makes for an interesting close-up account of infant OW production from a more longitudinal perspective. In Chapter 2 we saw how OWs matched the infant's preferred prosodic structures in the early stages of word production, and so to continue here with a wider and more in-depth analysis will enable us to analyse the nature of the output at different stages of language development: from the influence of babbling on the production of OW forms to the later phonetic motivations for the move towards CW production. Further data from Elsen (1991) will be drawn upon here to analyse the production of OWs, from the very first word until the 500-word point (the point at which Annalena has 500 word forms in her lexicon) at 1;6. This diary account provides a detailed picture of the progression of OWs, from the use of rudimentary vocal gestures to the eventual production of the CW form. The forms that Elsen describes include standard onomatopoeia that are commonly found in many languages, as well as highly idiosyncratic onomatopoeically-derived forms, or 'innovative OWs', which are far-removed from the adult language.

3.2 Methodology

3.2.1 Participant

Elsen (1991) provides a highly detailed chronological account of the language development of her daughter, Annalena, an only child acquiring German in Munich, Germany. Her observations begin shortly after Annalena's birth, and continue with a constant and thorough account of her language development until Annalena reaches 2;5. Elsen acknowledges the precocious nature of her daughter's speech, and reports that this was also noted by doctors; indeed, Annalena had already produced five different words by 0;9.

3.2.2 Data

Elsen's (1991) diary account was recorded on paper and included some audio back-up, though this was not available for this analysis. Exact dates are given and variability within word forms is highly detailed. Elsen also provides notes for most words, outlining whether they were produced spontaneously or whether they were elicited, and in which context the words appeared.

3.3 Analysis

Annalena's word production was analysed from the very first word until the 500-word point and OWs were identified in this dataset. A word was classed as onomatopoeic if it attempted to imitate a sound from the environment; this definition therefore includes commonly-recognised OWs such as *woof* or *vroom* as well as those which have been adopted or created by Annalena such as [lɑlɑ] for *Musik* 'music' and [bagbagba] to express *quak quak* 'quack quack', as discussed in Chapter 2. Forms carrying referential meaning in Annalena's output – 'functional' onomatopoeia, such as [vavaɯ] meaning *Hund* 'dog' – and those which are merely imitations of sounds from the environment (as in [bɯma] *bum* 'boom', to imitate a loud noise) will be considered in the initial analysis, before a full investigation of Annalena's functional OWs takes place. Naming conventions provided in Elsen's account will be adhered to here (e.g. [vavaɯ] appears under the target *Hund* 'dog', even though it is clearly derived from *woof* or *bow-wow*), and to avoid discrepancies in naming, the target word will be considered to be the full conventional form, and will appear in small caps in English (e.g. DOG). While in many cases the OW will remain in Annalena's lexicon as part of the adult language, this analysis considers the lexical and phonological transition from OW to CW, and so DOG relates to all word forms used to express *dog* throughout the recording period.

3.4 Results

3.4.1 Lexical acquisition

MOUSE [pr̥pr̥pr̥], at 0;10, is Annalena's very first OW, occurring 37 days after the onset of word production. From here, her lexicon increases rapidly in terms of both OWs and RWs. As we have already seen in Chapter 2, the most striking observation to be made regarding the presence of OWs in Annalena's output is the quantity of these forms that are acquired in the first months of word use. Figure 3.1 plots the monthly acquisition of OWs against that of RWs, where it is clear that OWs constitute an important portion of Annalena's total lexicon between 0;10 and 1;0, peaking at 0;11, at which point the acquisition of RWs starts to increase (see Figure 3.2). At 0;11 Annalena's lexicon can be seen to increase by 155%, with the acquisition of 11 new OWs. By 1;0 Annalena has 50 words in her lexicon, and the rate of OW acquisition remains constant at an average of just under six per month until 1;5. The acquisition of new OWs stops abruptly at 1;6 – the month in which Annalena reaches her 500-word point.

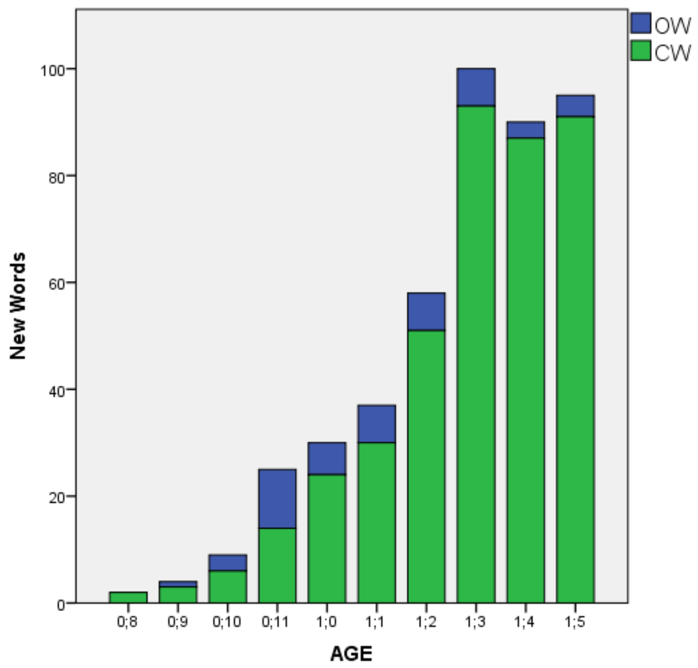

Figure 3.1: Monthly acquisition of OWs and CWs.

In Figure 3.2 the proportion of OWs in Annalena's output is plotted as a percentage of the cumulative total of new word forms (including OWs and CWs) acquired per month. The steady tailing-off of OW-acquisition over this period is equally apparent here.

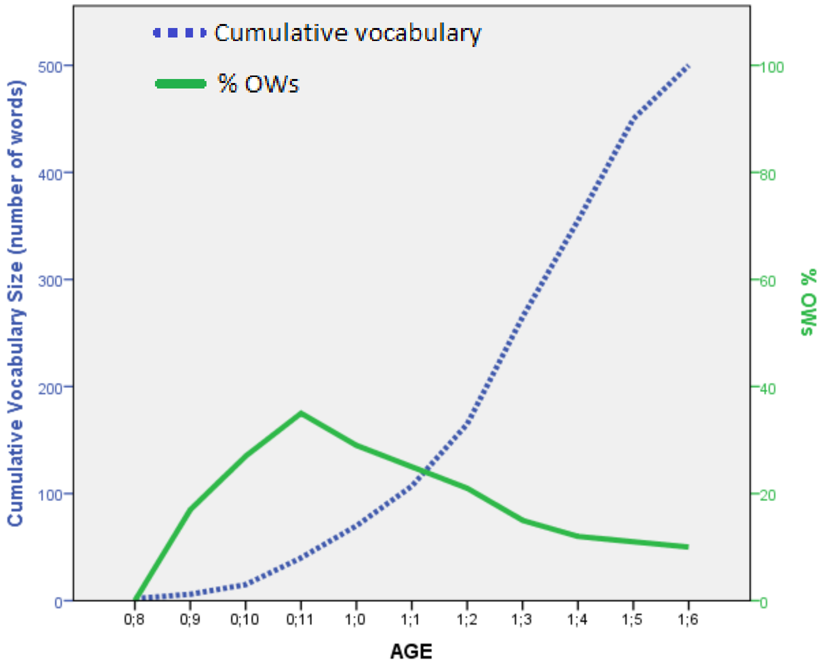

Figure 3.2: Percentage of OWs in relation to vocabulary size.

As can be seen in Figures 3.1 and 3.2, these forms initially account for a substantial portion of the total vocabulary (almost 40% at 0;11), before the acquisition of RWs takes over. From this analysis it is clear that the production of OWs contributes considerably to the early output, but once word learning takes off these forms no longer play an integral role in promoting word production. A closer look at the specific OWs in Annalena's output reveals some interesting features that are particular to these forms.

3.4.2 Referential onomatopoeia

Elsen makes clear in her account that many of Annalena's OWs serve as more than simply the verbal imitation of natural sounds. While some are produced only in this traditional sense (for example, *tatütata* 'nee-naw' (siren noise) [tɨtatɨta], *kikeriki* 'cock-a-doodle-doo' [kɪːkɪːç]), many OWs in the dataset are reported to be used referentially; that is, the production of OW forms in reference to objects, rather than just the sounds that those objects make (for example [mː] 'moo' to express *Kuh* 'cow'). We also see ample evidence of the creation of idiosyncratic OWs, which have already been referred to in Chapter 2, and which Elsen (1991, 1994) discusses in some detail in her own analyses. Annalena's use of *quak* 'quack', *piepiep* 'tweet' and her interpretation of a crow are declared to be unprompted and spontaneous on the infant's part; Annalena is reported to have imitated the sounds that these animals make and thus created her own lexical OW forms from these imitations. In reference to onomatopoeia, Elsen (1994) states that Annalena "imitated forms deliberately" (p.311), suggesting an infant-led approach to lexical development whereby OWs were used to facilitate production. Additionally, many of Annalena's OWs are produced with extra-phonetic features, which bring the sound and the meaning of these words closer together by making the lexicalised forms sound more realistic. This corresponds to Werner and Kaplan's (1963) discussion of onomatopoeia and how these forms function in the early output: Annalena's limited vocabulary does not prevent her from word production, as she creates her own protowords which serve communicative function based solely on the sound-meaning correspondences of these onomatopoeic creations.

OWs thus provide Annalena with a comprehensible sub-lexicon with which she can communicate and express meaning effectively, despite not having acquired the appropriate CWs. From the dataset of Annalena's first 500 words, 52 are judged as OWs, and 16 of these are used referentially. These are listed in Table 3.1 below, along with details of the extra-phonetic or prosodic features realized by Annalena in their production. An OW is classed as being referential if it serves to define more than merely onomatopoeic expression in Annalena's output – OWs which are adapted referentially as 'names for things' (e.g. [bagbagba] for DUCK).

Table 3.1: Referential OWs in Annalena's output.

Gloss	Target CW	Infant form OW	Extra Features	Age	Infant form CW	Age
BEE	Biene /binə/	[z̰ː]	"buzzing" sound	1;2	[bɪnə], [mɪːmə]	1;4
BIRD	Vogel /fɔgəl/	[pɪʳpɪʳpɪʳ]	Produced at a high pitch	0;11	[ʔəgl]	1;2
CAR	Auto /aɯto/	[ʙm], [bvːm]	Bilabial trill	0;11	[atɔ], [ɑːtɔ], [ɑːdɔ], [aɯtɔ], [atɔː]	1;2
CAT	Katze /katsə/	[mɛnːə], [mənːmɛɪnə]	High variability	0;11	[daksə], [daps], [kaça], [kax]	1;3
COW	Kuh /kuː/	[mː]	Long, single consonant	1;0	[gɯː]	1;4
CROW	Krähe /kʁeːə/	[bɔa]	Articulated with a hoarse voice (Elsen, 1996)	0;11	[kʁeːə]	2;1
DUCK	Ente /ɛntə/	[bag˺bag˺ba]	Unreleased plosives	0;11	[ʔɯnt]	1;2
DONKEY	Esel /ezəl/	[ˀiˀa]	Glottal closure + vowel combination	0;11	[ʔez̥ɛl]	1;5
DOG	Hund /hʊnt/	[bɑːvaɔ], [vavaɯ]	High variability, impersonation of dog	0;11	[ʔænðɛ], [anðə], [ænːtθə]	1;3
EAT (V)	Essen /ɛsən/	[namnam]		0;11	[ʔəsːa], [sːa], [aθːa], [ʔaθːa], [ɛsːə]	1;3
HARE	Hase /haːzə/	No transcription	"snuffling"	0;10	[ɑːzə], [ada], [baba], [hɑːdzə], [has]	1;3
HORSE	Pferd /feːɐ̯t/[8]	[ɨː]	"trembling" noise produced in falsetto	1;2	[feədə]	1;5
MOUSE	Maus /maɯs/	[pɪʳpɪʳpɪʳ]	Long vowels, high pitch, not fully articulated	0;10	[naɯ], [maɯsː], [vasː], [maɯbaɯðbaɯðmaɯ]	1;3
MUSIC	Musik /muzik/	[lɑːlɑː]	Often sung	1;1	[çɪk], [zɪːkʰ]	1;5
PIG	Schwein /ʃvaɪn/	No transcription	"grunting"	n.s.	[waɪ], [baɪ], [vaɪn]	1;4
SHEEP	Schaf /ʃaːf/	[mɪː], [meː]	Vowel variability	1;1	[ʃʲaɔf]	1;8

[8] Standard pronunciation of *Pferd* would be [pfeːɐ̯t]; the transcription in the table reflects the mother's dialect.

As is apparent from Table 3.1, these forms include a range of extra-phonetic, or 'wild', features (Rhodes, 1994), produced in imitation of the sound representing the word in question. These wild features include segments that are not found in the native language phonology, unusual pitch and intonation changes, gestural accompaniments, variability, and the unconventional production of familiar phonemes. The examples in Table 3.1 illustrate the idiosyncratic nature of OWs in Annalena's output, and the various ways in which meaning can be conveyed through the use of sound effects in their production, despite these forms being far removed from the phonological conventions of the adult language.

With regard to infants' referential OW production, it is reported that these forms "generally disappear quickly in favour of the traditional forms" (Jespersen, 1922, p.151). However, evidence from Chapter 2 shows that this is not always the case, and Annalena produces her referential OWs for some months before acquiring the full adult form. It could be assumed that Annalena would acquire CWs over time as she develops a more complete phonological inventory, but as Table 3.1 shows, she continues with the OW production of SHEEP until age 1;8, and produces OW CROW until after her second birthday. It is evident that a complex system of language learning is at work here, and this will be explored in detail in this analysis.

Figure 3.3 provides a detailed depiction of the transition from OWs to CWs in Annalena's output over time, as initially shown in Figure 2.3 (Chapter 2). We have already observed how OWs appear well before the CW is first produced in all instances, but in this case we go beyond the 100-word point, further highlighting the long delay between OW acquisition and eventual CW production. Figure 3.3 reflects the trajectory of lexical development as OWs are initially used referentially, representing the CW equivalent that Annalena will eventually move towards in production. Here we also see that many forms undergo a period of alternation, as Annalena produces both the OW and the CW in tandem to express the same meaning, eventually moving towards the CW; at this point the OW no longer serves to carry referential meaning. During this transitional period Annalena produces both individual tokens of OWs and CWs (for example, at 1;0 she produces DUCK as both [bagba] and [nɛndə]), as well as combinations of the two, as in COW [mːkɯç] at 1;5. This transitional phase marks an incremental shift in representation from the OW to the CW form, and once again returns us to Werner and Kaplan's (1963) discussion of OWs in early word production, possibly reflecting a period of increasing differentiation as infants move from the minimally-differentiated OW forms to the arbitrary sound-meaning correspondences of CWs.

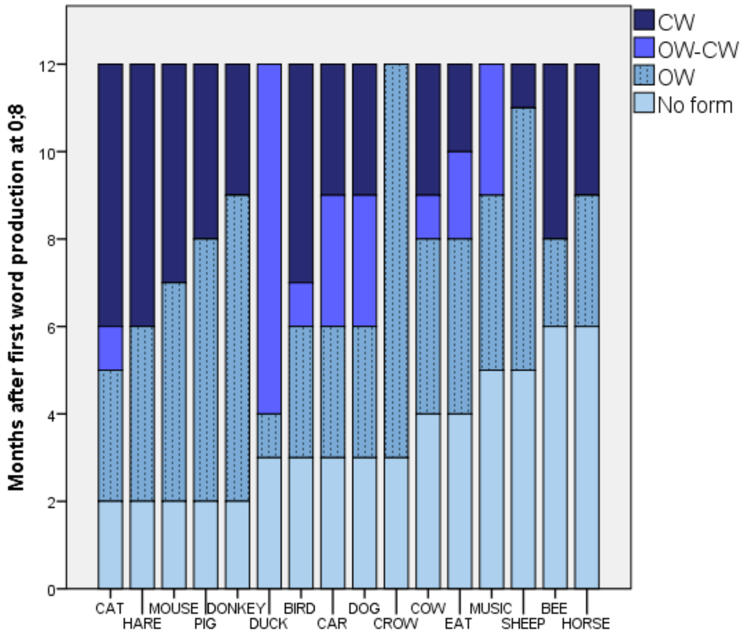

Figure 3.3: Progression of Annalena's OWs over time.

The changes taking place in Annalena's wider output can account for the transitional phase observed here, as we see consistencies over time in the nature of these alternating forms. Fifty percent of OWs alternate with the CW in a transitional phase before the CW fully takes over. With the exception of CAT and CROW, all of the forms which undergo a period of transition are acquired between three and five months after first word production: this corresponds to the period of 0;11 to 1;1, just before lexical acquisition becomes exponential at 1;2 (see Figure 3.2). It is unsurprising that we find evidence of shifting representations here, as Annalena's lexical knowledge begins to increase rapidly as new word forms are acquired, and thus new forms are likely to be less stably represented in the output. Those forms which do not undergo a period of alternation are mostly acquired either in the earliest stages of word learning, and are thus more stably represented in Annalena's lexicon, or are acquired later, when Annalena's production experience is much more established. It seems that this period of transition is not limited to her OWs, but is taking place across the board in her language development: new linguistic knowledge is being acquired, causing Annalena's lexical and phonological systems to become less stable in terms of both internal representation and output.

Fundamentally, the transition from OW to CW does not take place across the board but on a word-by-word basis. More than two thirds (69%) of her OWs make the transition to CW after 1;2 – coinciding with the steep increase in lexical knowledge shown in Figures 3.1 and 3.2. Changes taking place elsewhere in Annalena's phonological development can account for these lexical advances; the following analysis provides evidence towards a phonological shift in Annalena's language development, which coincides with the period of transition noted in Figure 3.3, paving the way for the lexical advances shown in Figures 3.1 and 3.2.

3.4.3 OWs and segmental development

An important development takes place in the first half of Annalena's second year which enables the change in representation from OW to CW for many of the forms in Table 3.1. Phonologically-derived motivations for the segmental differences between the OWs and their CW counterparts can be identified when Annalena's full output is examined; the developments in her consonant inventory between 1;0 and 1;2 (Figures 3.4a and b) show a shift in her segmental capacity which takes place shortly after her first birthday.

	Bilabial		Dental		Alveolar		Postalveolar		Palatal		Velar		Uvular	Glottal
Plosive	p	**b**			**t**	**d**					k	**g**		**(ʔ)**
Nasal	**m**				**n**						**ŋ**			
Trill	(ʙ)													
Fricative	(β)		f	v	s	z	ʃ	ʒ	ç		x		**ʁ**	**h**
Affricate			pf		ts		tʃ	dʒ						
Approx.	(w)					**l**				**j**				

Consonants produced in at least two word tokens are highlighted in blue. Consonants produced only in OWs are highlighted in green. Consonants which are not highlighted have not yet been acquired. Brackets denote phonemes which do not appear in Standard German.

Figure 3.4a: Annalena's consonant inventory at 1;0.

	Bilabial		Dental		Alveolar		Postalveolar		Palatal		Velar		Uvular	Glottal
Plosive	**p**	**b**			**t**	**d**					**k**	**g**		**(ʔ)**
Nasal	**m**				**n**						**ŋ**			
Trill	(ʙ)													
Fricative	(β)		**f**	**v**	**s**	**z**	ʃ	ʒ	**ç**		**x**		**ʁ**	**h**
Affricate			pf		ts		tʃ	**dʒ**						
Approx.	**(w)**					**l**				**j**				

Figure 3.4b: Annalena's consonant inventory at 1;2.

Those segments which had appeared only in OWs at 1;0 have stabilised in CWs by 1;2, while Annalena has acquired almost a full set of fricatives during this period. The changes seen across these two consonant inventories can account in part for the shift from OWs to CWs that begins at 1;2.

Figure 3.4a shows how certain segments were produced exclusively in OWs, reflecting the unconventional characteristics of these forms which do not necessarily match the phonological framework of the ambient language. Before 1;0 Annalena produces the bilabial trill /ʙ/ exclusively in CAR (see Table 3.1), and the approximate /w/ only in DOG (see Table 3.1): neither are typical phonemes of German. A predominance of labial consonants is reported in Annalena's babbling phase, including /w/, /b/, /bv/ and /ʙ/, which accounts for the continuing use of these segments in OWs; these segments enable Annalena to represent words using the well-rehearsed phonemes that are available to her. These examples demonstrate how OWs may work functionally in early word production; despite Annalena's inability to consistently produce the segments necessary for the established German OW forms (*brumm* and *wauwau*, respectively), her own forms can still be understood in communication with the caregiver. Furthermore, the use of non-native phonemes in these OWs appears to facilitate further word production, as Annalena later goes on to use these segments in the production of RWs by adapting word forms with these consonants: /w/ is produced in place of fricatives in a number of words, including *Affe* 'ape' [ɔwa], *zwei* 'two' [waɪ] and *Vorsicht* 'look out' [was:ɪ:ç], as well as in place of word-medial plosives in *Auge* 'eye' [awa] and *Opa* 'grandfather' [ɔwa]. These early consonants facilitate word production over time, as Annalena practices and stabilises them with the production of OWs, and later acquires new words by adapting them to fit her own consonant inventory.

As well as these non-native phonemes, Annalena produces a number of German segments in OWs which are avoided in CWs: /k/ is produced in *kikeriki* 'cock-a-doodle-doo' [kɪːkɪːç] and *killekille* 'tickle tickle' [kɪɫ'əkɪɫ'ə] but is avoided in *kalt* 'cold' [malː]; /ç/ also appears in *kikeriki* whilst being dropped from the cluster in *Milch* 'milk' [malː]; finally, /v/ is produced in the OWs DOG and CAR at 0;11 (see Table 3.1), but is not used consistently in a RW until 1;4 with *wischen* 'wipe' [vɪʂ̥ːə, vɪʂ̥ːə'n]. Again, all of these consonants are reported to first occur in Annalena's babbling phase, but here they are limited to OW forms, preventing more accurate or varied word production at this point.

3.4.4 OWs and phonological development

We can also see consistent trends throughout Annalena's OWs which point to a phonological role for these forms, leading on from findings observed in Chapter 2. In a paper investigating the phonological constraints in Annalena's output, Elsen (1994) describes her daughter's tendency to avoid certain forms, including a refusal to produce the word *Hund* 'dog', despite being presented with plenty of opportunities to imitate this form from those around her, and despite a clear understanding of the CW. Elsen goes on to discuss Annalena's approach to word production in terms of the 'principle of avoidance', as the infant is reluctant "to produce sound combinations beyond her phonological level" (1994, p.311). This account suggests that a phonological limitation is prompting Annalena's use of OW DOG, which enables communication while her phonological skills catch up with her wish to express certain word forms. Ferguson and Farwell (1975) describe infants' process of avoiding specific words and phonemes as one of "great selectivity" (p.433), bringing us back to the previous analysis in Chapter 2 and the selection of OWs over their CW counterparts based on an infant's preferred prosodic structures. In this case, Annalena can be seen to select the OW *wauwau* over CW *Hund* to correspond to the well-rehearsed prosodic patterns from her output which are more suited to the OW form. By doing so she is matching her own vocal capabilities to specific lexical items.

A thorough phonological analysis of the changes taking place in this period can further demonstrate how Annalena makes use of OWs to overcome the phonological limitations presented by the CW forms. This can be seen across her lexicon, and as well as the prosodic matching of these forms that we observed in Chapter 2 we also see a role for Annalena's developing articulatory capacity: the segmental changes observed between Figures 3.4a and b mark a phonological shift in the early output, prompting a transition to the CW forms as well as the acquisition of many new words. An account of the phonological limitations in Annalena's output in relation to her functional OWs demonstrates this point.

Fricatives

As shown in Figure 3.4a, fricatives are highly restricted in Annalena's early words, and in fact at 1;0 fricative segments can only be found in OWs. They begin to make their way into Annalena's output between 1;1 and 1;2, both through the acquisition of new words and through the transition from OW to CW. Of the 24 target words containing fricatives that were attempted at 1;1, only 25% were realised with a fricative, while at 1;2 64% were attempted with a fricative (compare *Kissen* 'cusion' at 1;1, [gɛjə], and 1;3, [kɪs:], [kɪsʃə], [kɪːça], by which time /k/ had also been acquired – see below). This highlights a change in Annalena's approach to word production over this short timespan – the shift at 1;2 finally sees Annalena attempting to produce (and often succeeding in producing) fricatives in CWs, and in turn this leads to the transition of many OWs to the CW equivalent.

BIRD (*Vogel* /fɔgəl/) presents the first example of fricative avoidance in Annalena's OWs. Word initial /f/ is not reported to stabilise until 1;5, and we don't see any correct attempts at producing this segment until 1;3, at which point CW *Vogel* is attempted as [vɯgl̩] (with high variability), alongside *vier* 'four' [fɪa, vɪa], *Fuchs* 'fox' [vɯg, ɯkz] and *Fuß* 'foot' [vɯs:]. In total, 19 new /f/-initial words are first attempted at 1;3, amongst which 14 include a word initial [f] or [v] in at least one token.

CAT (*Katze* /katsə/), DONKEY (*Esel*/ezəl/), EAT (*Essen* /ɛsən/), HARE (*Hase* /ha:zə/) and MUSIC (*Musik* /muzik/) all contain a word-medial /s/ or /z/ in the CW, and Elsen refers to the principle of avoidance in terms of these segments. Indeed, aside from two exceptions she claims that Annalena 'ignores' /z/ until 1;3 (1991, p.84, own translation), and accordingly, Annalena transitions from OW to CW in all five of these forms after 1;3. Word-final /s/ in CW MOUSE (*Maus* /maɯs/) exhibits similar restrictions, and Annalena acquires this form alongside the others at 1;3, with additional /s/-final words appearing in her lexicon during the same period (e.g. *Bus* 'bus' [bɯθ, bɯs̪], *Netz* 'web' [dɛts:]). However, we do see final /s/ as early as 1;1 in *heiß* ‘hot’ [ais] and *tschüß* ‘bye’ [tɨs̺, tsɨs], though Elsen reports an unstable production of this segment, which is not truly mastered until 1;5.

The initial /ʃ/ in SHEEP (*Shaf* /ʃa:f/) is highly problematic for Annalena, and she can be found to avoid this segment throughout her lexicon. Indeed, she does not produce word-initial /ʃ/ correctly until 1;5 with *schau mal* ‘look’ [ʃçaɯmaˡ], but even then this is one correct attempt amongst a range of variable tokens. Out of the 45 /ʃ/-initial words that are attempted by Annalena in her first 18 months, only five forms contain a /ʃ/ segment, and including *schau mal* only two of these are realised in the correct position – this segment is highly limited in Annalena’s output, and we do not see a correct attempt at SHEEP until 1;8 (although we must also bear in mind the word-final /f/ here).

Finally, initial /ʃv/ in PIG (*Schwein* /ʃvaɪn/) presents an especially difficult problem for Annalena, and this cluster is not mastered in her output until 1;9, which is beyond the timeframe of this analysis. However, owing to the acquisition of fricatives, Annalena is able to transition to the CW here as early as 1;5, as she makes use of cluster reduction in order to facilitate production, as in [vaɪn].

/h/

/h/ only occurs at morpheme-initial boundaries in German, and before 1;1 Annalena has attempted just two /h/-initial words, though she produces this segment correctly in certain tokens of both: [hal:, halɔ] *hallo* ‘hello’ at 1;0, and [haɪ, haɪç] *heiß* ‘hot’ at 0;11.

Between 1;1 and 1;2 she acquires 11 /h/-initial words, of which 8 are realised with an accurate /h/. This includes the CW form of HARE (*Hase* /ha:zə/), which is first produced at 1;2.

Word-initial /k/

Until the end of 1;1 Annalena had only produced /k/ in OWs. Its distribution is highly limited, and again Elsen (1991) refers to Annalena’s approach here as another example of the principle of avoidance. Before 1;1 Annalena attempts only one /k/-initial word, the OW form *kikeriki* ‘cock-a-doodle-doo’. At 1;2 Annalena begins to produce initial-/k/ correctly, in *Kater* ‘tomcat’ [kɑːtə] and *Kuckuck* ‘peek-a-boo’ [kɯkɯː], and acquires 10 new /k/-initial words between 1;2 and 1;3, of which eight have a correctly-produced word-initial segment in at least one token.

CAT (*KATZE* /katsə/) and COW (*Kuh* /kuː/) (see also CROW, below) are restricted by this phonological limitation in Annalena’s output, and accordingly neither of these OWs transition to the CW form until word-initial /k/ has stabilised in Annalena’s output, between 1;3 and 1;4.

Word-final /k/

/k/ is limited word-finally as well as word-initially, and is not attempted in this position until Annalena produces *Tag* ‘day’ as [tɑːkʰ, gɑːkʰ] (first produced at 0;11 as [dada]) at 1;2. Two more /k/-final words are acquired in the same period: *weg* ‘way, away’ [bakʰ] and the OW *ticktack* ‘ticktock’ [kəkək]. A few months later, at 1;5, the transition of MUSIC from OW to CW [zɪːkʰ] takes place.

Word-medial /t/

/t/ appears in *Teddy* at 0;10, but *bitte* ‘please’ is the only target word with medial /t/ in Annalena’s output in this period, produced with reduplication as [bɪtˑəbɪt] at 0;10. Otherwise, the first target word with this segment is *Kater* ‘tomcat’ [tɑːtə, kɑːtə] at 1;1.

Annalena’s first CW attempt at CAR (*Auto* /aɯto/)[ɑːtɔ] follows at 1;2, and in the same period eight other new word forms with medial /t/ are acquired, including *Tomate* ‘tomato’ [bɔmɑːtə] (1;2), and *Schmetterling* ‘butterfly’ [mɛtːtinː] and *warte* ‘wait’ [watˑa] at 1;3.

Word-medial /n/

There appears to be some restriction on the production of /n/ in Annalena’s output, though this is surprising as *nein* ‘no’ [nan] appeared as her first word at 0;8. However, while this segment is stable word-initially, Annalena has problems producing it word-finally, for example *Bein* 'leg' [baɪ] and *Stein* 'stone' [daɪ, daja] at 1;5. There are exceptions to this restriction, as she can produce /n/ in *an* 'on' [ʔan] at 1;3, though Elsen acknowledges this anomaly herself, positing that it may be due to the lack of any other consonants in this form.

There are few examples of word-medial /n/ in Annalena's dataset, which can often be found to exhibit consonant harmony: *Kanne* 'can' [nana] at 1;3 (with word-intial /k/ also presenting a problem here), *Annalena* [nana] from 0;10 until 1;5. *Banane* 'banana' is an exception, produced as [maɪna, mana] from as early as 0;11, but there is especially high variability in the realisation of this word.

BEE (*Biene* /binə/) transitions from the OW to the CW form [bɪnə] at 1;4, and we see the acquisition of a number of words with medial /n/ in the same period: *Honig* 'honey' [hoːnɪs̪], *Pony* [pɔnɪ] and *Spinne* 'spider' [pɪn˙ə] at 1;4, and *Birne* 'pear' [bɛn˙ɑːt] and *Sonne* 'sun' [z̥ɔn˙ə] at 1;5.

Cluster /nt/

Annalena avoids this cluster almost completely, with only one target word containing /nt/ (*Mund* 'mouth', produced as [mama, maɪ, manθ]) until its first correct realisation at 1;2. CW DOG (*Hund* /hʊnt/) [ʔɯnt] is one of three new words that Annalena produces with this cluster in this month, as well as *Kind* 'child' [tɪntʰ] and *Elefant* 'elephant' [man tʰ]. *Mund* is now also produced with the correct cluster, as [mɔntʰ], and at 1;3 we see the transition of DUCK (*Ente* /ɛntə/)from OW to CW with word-medial /nt/ [æn:tθə, ɛndə].

Cluster /kʁ/

It is unsurprising that the cluster /kʁ/, as well as many other clusters, is limited in Annalena's output. Elsen (1991) reports that this cluster is not acquired until 1;9, which goes beyond the period analysed here, but even so CROW (*Krähe* /kʁeːə/) is not fully acquired until 2;1, when it is produced accurately as [kʁeˑə].

3.4.5 Summary

This analysis has demonstrated a number of phonological limitations which affect the entirety of Annalena's output, but which can largely be seen to disappear after the period of transition between 1;0 and 1;2. Before these changes take place we can see that Annalena's OW forms provide, in almost all cases, a better 'match' to her phonological abilities than would the CW forms. This demonstrates the presence of a lexical alternative which can be systematically implemented whenever an OW is available as an equivalent to a challenging word form. This lexical alternative makes use of OW forms to facilitate production when the phonological skills necessary to produce the CW are not yet in place. Over time we see a general transition from OW to CW production in Annalena's data, as her phonological systems fall into place and she is able to produce the full CW forms. Alongside these transitions we see extensive lexical acquisition, prompted by each of the phonological developments discussed above. In many cases Annalena can be seen to avoid words containing these phonological segments completely, only attempting them once she has the appropriate capacity to do so. Her careful approach to CW production is thus applied throughout her lexicon, but in cases where there is no suitable OW to use referentially she must instead wait until her output capacity is sufficiently developed to produce these words.

In some cases these limitations affect Annalena's OW production as well as her CW production. OW BEE, for example, is produced as a single fricative [z̺ː], but is acquired at 1;2, once she has begun to acquire fricatives. Drawing again upon the example of BEE, after only two months Annalena is able to produce word-medial /n/, and thus she transitions to the CW form *Biene*, repeatedly moving forwards within the boundaries of her phonological capacity.

It must also be noted that many of the CWs in this analysis present multiple challenges for Annalena: CAT, DOG, HARE, MUSIC and SHEEP all contain two of the phonological limitations listed above (see Table 3.2). Annalena does not attempt to tackle both segments simultaneously, but typical to the systematic approach that we observe throughout her production, she adapts CW forms in line with her expanding production capacity, moving gradually towards a more accurate production. The gradual transition to these five CWs is shown in Table 3.2. We can see that the forms listed under CW1 take on only one of the newly-acquired phonological segments, while under CW2 we see a gradual transition towards the second phonological segment, with instability across all forms (and regression in SHEEP) as she does so. By CW3 she is close to the full adult form in all cases.

Table 3.2: Transition to CW forms over time.

Gloss		CW 1	CW 2	CW 3
CAT	*Katze*	[aka] (1;2)	[taθːə] (1;3)	[kaxːa] (1;3), [daksə, kaça] (1;4)
DOG	*Hund*	[ʔɯnt] (1;2)	[hɯtʰ] (1;4)	[hɯnt] (1;6)
HARE	*Hase*	[ɑːzə] (1;2)	[ɔsːɔ, hɑːsə] (1;3)	[hɑːzə] (1;3)
MUSIC	*Musik*	[ʲɪk] (1;5)	[zɪːkʰ] (1;5)	[mɯɯsɪːk] (1;10)
SHEEP	*Schaf*	[ʃʲaf] (1;7)	[ðɑːf] (1;8)	[z̠ɑːfə, ʃʲaf] (1;9)

Furthermore, the acquisition of these forms is not delayed by the extra phonological challenge that they pose; three of the five are acquired as early as 1;2, corresponding in all cases to the timeframe in which these phonological changes start to take hold, and with the transition to CWs across Annalena's OW forms.

3.5 Discussion

The production of OWs is prominent in Annalena's early output and contributes substantially to her early vocabulary. This prominence does not last, but leads to a period of rapid phonological change in the output – from 1;0 to 1;2 – which prompts lexical development across Annalena's output. OWs have been shown to be highly productive during this period: we see the development of the consonant inventory as well as a steep increase in her lexical and phonological capabilities. As new segments are acquired, changes are observed across the board, with lexical acquisition becoming exponential as new articulatory skills are refined and stabilised. Annalena's output appears to be in transition, yet her use of OWs – acquired over the first months of word production – remains stable until these changes are established, after which point a general move towards the production of CW forms takes place. This demonstrates the systematic use of OWs in Annalena's lexicon, as these forms enable her to maintain stability in an increasingly unstable output.

Early on, we see a segmental advantage for OWs, as consonants that were rehearsed in Annalena's babbling phase are used exclusively in the production of these forms. These are not necessarily typical phonemes of German, yet they provide an inventory of segments which are appropriate for imitating sounds in her environment (the bilabial trill in CAR, for example, or the labial fricative in DOG), and thus allow Annalena to begin forming her early phonological system. This corresponds to Jakobson's (1980) claims, whereby the phonetic flexibility of OW forms is considered to provide motoric and articulatory refinement in early vocal production. In Annalena's early data, OWs are seen to facilitate production through the mastering of a wider variety of consonants, which later lead to the expansion of the lexicon as these consonants – despite some being atypical of the ambient language – are used productively in the adaptation of new words.

When considering Annalena's early consonant inventory (Figure 3.4a), one might assume that she should be able to produce CWs COW, BEE, CROW, DUCK, DOG and HORSE, as all of these forms contain segments that have already occurred in her output. However, while her articulatory capacity is relatively developed even by 1;0 – indeed, Annalena is already able to produce over 70 word forms at this point – phonologically-speaking, her output is heavily restricted. It is here that we see the implementation of OWs on a phonological level, offering an alternative form which can be used to communicate referential meaning. In most cases OWs match her phonological ability – OW BEE, for example, appears at 1;2, once Annalena is able to produce fricatives, while others have simple CV or consonant harmony structures as observed in Chapter 2. However, we do see examples of OWs which override these limitations – fricatives are found in OW DOG and CAR, and medial /n/ in CAT – but these apparent exceptions to Annalena's rigid phonological rules should be considered in terms of their initial appearance in her output and Annalena's wider trajectory of phonological development. All three of these forms were first produced at 0;11, before Annalena had 40 words in her lexicon; OWs DOG and CAR were acquired amongst her first 25 words. Here we can see examples of early words being selected on a whole-word basis to fit the prosodic structures that are most common in the early output, as was shown in the previous chapter. Only later do we see evidence towards a phonological reorganisation – reflected in the changes observed between Figures 3.4a and b – at which point lexical acquisition becomes dependent on Annalena's segmental, rather than prosodic, capacities. Annalena has ample experience of these forms owing to their early acquisition, and is able to continue producing segments in these contexts which are otherwise avoided in her expanding lexicon. Here we see examples of 'frozen forms', which go against the systematic trends in her output owing to their early acquisition.

Elsen (1994) considers the principle of avoidance in a paper exploring the lexical and phonological constraints in Annalena's output, describing how Annalena produced "simple forms…which enabled [her] to speak of corresponding referents successfully" (p.313). While to some extent her argument supports observations made in this analysis, it seems that Annalena's approach is more than simply the *ad hoc* avoidance or substitution of specific words and phonemes. Indeed, her approach is highly methodical, and rather than avoiding the production of certain words altogether, Annalena applies an OW alternative which enables her to gain valuable production practice within the limits of her phonological system. The systematicity of Annalena's OW use is striking, with a pattern force which brings about the creation of new OWs, leaving no motivation for the acquisition of the equivalent CW. A similarly systematic approach can be seen in later word production, as she acquires the typical CW equivalents across the board, bringing together her phonological and lexical abilities to reflect a more adult-like production.

From Elsen's own accounts (1991, 1994) we have a wealth of evidence showing that many of Annalena's OWs – BIRD, CROW, DUCK, HARE, HORSE and PIG; or 38% of her referential OWs – are not based on lexicalised OWs from the input, and must thus have been adapted by Annalena to replace challenging CW forms which lie outside of her phonological capacity. This returns us to Werner and Kaplan's (1963) theoretical viewpoint, which could provide an appropriate interpretation of this data. Indeed, we see numerous examples of Annalena's own creation of OWs, using these forms as labels for abstract concepts in her surrounding environment. We also see what could be considered as a period of increasing differentiation, when Annalena alternates OW and CW forms (and sometimes combines them, as in [mːkɯç] *muh-Kuh* 'moo-cow'), supposedly reflecting an increasing ability to abstract forms and meanings in language development. However, this analysis leads to the conclusion that it is the segmental and phonological features of OWs which appear to have the most prominent role in the early acquisition of these forms. It has been suggested that onomatopoeia may be easier to produce in early development (Kunnari, 2002), and indeed the OWs observed here have been found to exhibit features that do not necessarily match those of the language being acquired, or the expected output constraints that are implemented elsewhere in the lexicon. In most cases, the 'articulatory shape' of these OWs is variable owing to the marginal nature of these words; their fuzzy phonological boundaries appear to facilitate production by providing an impressionistic template which incorporates lexical meaning while also allowing a wide margin of phonological error. Indeed, many common characteristics of OWs – prosodic and extra-phonetic modifications, reduplication, vowel lengthening – are likely to require less planning, and demand less precise motor skills in production; it is therefore unsurprising that they feature so prominently in Annalena's early output. In addition, these common features appear to enable the development of phonological precision through the production of word forms which are motorically uncoordinated and variable, but which can still be understood and are thus productive in communication.

A case study of only one infant cannot provide a broad perspective on the use of OWs, but from this analysis we have evidence to suggest that OWs do indeed play an important role in at least some infants' language acquisition. Here we have found OWs to fit within Annalena's individual linguistic system, not as outliers in any phonetic, phonological or even semantic sense, but as usable meaningful units which enable communication with the caregiver, while also promoting production and thus the development of articulatory and planning capacities. In Annalena's output at least, these forms are found to play a useful and relevant role in the trajectory of her language development by simultaneously being both idiosyncratic to and typical of her general output. In Chapter 2 we observed how OWs featured differently across nine infants' outputs – Annalena had one of the highest proportions of OWs across the dataset, while some infants produced only a couple of these forms – and an analysis of any of these other eight infants would no doubt provide an entirely different set of results. However, we did observe consistencies across the infants in terms of the prosodic structure of OWs, and in this way we can assume that some trends may also exist in the phonological detail of early OW acquisition. Indeed, just as we observe universals and idiosyncrasies in language development as a whole, so too can we expect to find trends or even consistencies in infants' OW production. The main limitation of this study is that it is not possible to account for these trends without access to datasets similar to that provided by Elsen (1991), which gives ample phonetic detail and variability, as well as a longitudinal record of Annalena's entire vocabulary. However, a consideration of this analysis alongside findings from Chapter 2 allows us to speculate on the 'role' of OWs in early development. It seems that these forms may not have a specific role *per se*, but may simply fit neatly within the infant's expanding and evolving early productive system.

3.6 Conclusion

OWs have been shown to provide a lexical alternative which allows Annalena to bridge a gap in her production capacity. As shown in Chapter 2, OWs are 'selected' in early production for their simple prosodic structures which match her early output capacities. In all cases CW forms presented phonological challenges beyond Annalena's ability, but here we observed the eventual transition to these forms as she moved from a whole-word to a segmental approach in her word production. We have thus observed how OWs may serve as a bootstrapping mechanism for word learning, as OWs provide a second option to some of the rigidly-structured adult forms towards which Annalena eventually moves. Selection on both a phonological and a lexical level is apparent as she matches the words she produces with her phonological capacity.

From this case study it can be posited that OWs are more than just an irrelevant feature of a temporary babytalk lexicon. OWs were found to be abundant in the output, working alongside the developing lexicon rather than existing as a subsidiary feature to CWs. In Annalena's case at least, onomatopoeia have been found to be lexically meaningful while also serving productively at the front line of phonological development.

SECTION II

Infants' perception of onomatopoeia: two eye-tracking studies

In Chapters 2 and 3 we observed a match between infants' production of OWs and the general prosodic and phonological features that were identified amongst their earliest words, suggesting a usage-based motivation for the acquisition of these forms in the early stages of lexical development. However, much of the sound symbolism research demonstrates infants' sensitivity to sound-meaning correspondences even before they are able to speak (Asano et al., 2014; Ozturk et al., 2013). Indeed, representations of both OWs and CWs are formed through experience in the input long before infants are able to produce them, and so we must consider whether the iconicity of OWs may provide a perceptual advantage over CWs in the establishment of early form-meaning mappings. Bergelson and Swingley (2012) found that infants as young as 6 months could match pictures to words when spoken by their mothers, although performance was noted to be much more reliable after the age of 14 months. Furthermore, infants as young as 11 months have been shown to recognise familiar words produced by an experimenter, although contrasting findings have been observed with regard to whether these representations are phonologically underspecified (Hallé & de Boysson-Bardies, 1996) or finely detailed (Swingley, 2005). Friedrich and Friederici (2011) used a word-learning task to show that 6-month-olds were able to learn novel word-object pairings after relatively few exposures, and these pairings were still familiar to the infants in a memory task carried out one day later. Additionally, Friedrich (2008) reports that infants are able to distinguish between congruous and incongruous word-object pairings by 12 months of age. However, even at 12 months these infants failed to demonstrate more complex semantic integration processes in the incongruous condition (unlike infants of 14 and 19 months in the same experiment), which suggests that infants in this younger group were only just beginning to establish form-meaning mappings. Together these findings reveal some uncertainty as to precisely when infants begin reliably mapping word-meaning pairings in early perception, and indeed developments in the output are known to be strongly correlated with infants' semantic integration abilities (Friedrich, 2008; Werker et al., 2002).

With these findings in mind, we must consider whether OWs may be more easily represented in memory in early language perception. Imai and Kita's (2014) 'sound symbolism bootstrapping hypothesis' posits that young infants' sensitivity to sound symbolism provides them with a 'referential insight' (p.2) into the sound-meaning correspondences in onomatopoeic words, thus establishing a lexical representation for these forms more easily, and perhaps earlier, than their conventional counterparts. This would support data from Chapters 2 and 3 showing that these forms are often found in an infant's early lexicon prior to the CWs (see Figure 2.3) – a finding that also corresponds to Werner and Kaplan's (1963) theoretical position.

The following analyses will consider young infants' perception of OWs in order to investigate whether OWs have an advantage over CWs in early perception. Here we will account for infants' experience of OWs in real time, addressing the question of iconicity in order to identify any role that their less-than-arbitrary sound-meaning links may play in the processing of these forms. Two paradigms will be explored in the analyses that follow: one accounting for the extent of iconicity in the OW itself ('wild' versus 'tame'; Rhodes, 1994) and one comparing responses to the presence or lack of iconicity in word forms overall (OW versus CW). This will enable us to understand whether perception is indeed facilitated by the sound-meaning links that are posited as being advantageous to infants in early language learning.

Eye-tracking will be used in the two analyses that follow to trace infants' mappings of auditory and visual stimuli in real-time. This is just one of a number of methodologies which have recently begun to gain ground in developmental research: eye movements, as well as changes in heart rate (Colombo et al., 2001), pupil size (Laeng et al., 2012) and event-related potentials (ERPs; Asano et al., 2015), provide a measure of infants' cognitive processing of simultaneously-presented auditory and visual stimuli, providing a multimodal view of early perception. Indeed, eye-tracking has become "an increasingly important tool in research on visual perception" (Hayhoe, 2004, p.273), and has been used to successfully generate novel findings throughout the field of developmental psycholinguistics, and across a range of populations including newborns (Turati et al., 2010) and infants at risk from autism (Merin et al., 2007).

The difficulties of working with eye-tracking, especially with infant participants, are readily acknowledged in the literature; Aslin (2007) even claims that the issues encountered in eye-tracking research may suffice as being "motivation to seek employment in another scientific discipline" (p.5-6). Oakes (2012) discusses the problematic programming issues involved in combining common eye-tracking systems (the analysis capabilities of which are often too limited for use in infant research) with more sophisticated software packages such as Matlab and E-Prime, whereby the use of multiple packages can lead to a loss of accuracy in the analysis, as visual and audio stimuli lose synchronization through delays across systems (see Aslin, 2012, Appendix, for an assessment of this problem). Alongside these technical issues, which will no doubt cease to be problematic as understanding in this domain increases, the analysis of the data also presents a notable challenge. Fixations – that is, holding the pupils' focus on a single location for even a brief moment – can give us differing accounts of perception across participants, as readings can vary depending on the age of the infant and the developmental stage that they may be at (Oakes, 2010). Furthermore, the multifaceted nature of fixation data – which allows us to measure location and duration of fixations, response time, accuracy and anticipation, among other features (Hayhoe, 2004) – can make it difficult to explore specific research questions, and in this respect, different measures of the data may yield opposing results. Furthermore, Aslin (2012) points out that any findings are based on the assumption that the direction of an infant's gaze is directly associated with their cognitive processing at that very moment, when in fact many differing variables may contribute to the gaze-patterns exhibited by an infant when presented with a certain set of stimuli. Aslin (2007) also stipulates that, despite the large amount of data that an eye-tracker generates over the course of any experiment, still a lot of potentially important information is disregarded, leaving only a global measure of infants' fixations and eye movements over the course of the procedure.

Despite these potential drawbacks, eye-tracking remains a useful tool in the understanding of infants' linguistic and cognitive processing, and can be used to analyse a range of perceptual measures, including discrimination, preference, categorization, detection, learning and expectation (Aslin, 2007). Furthermore, a broader knowledge of infants' responses in relation to their eye-movements is leading to more sophisticated analyses in this domain, as researchers come up with novel ways of using the extensive data produced during eye-tracking experiments to explore new methodologies in this field (Aslin, 2012).

In the two experiments that follow, eye-tracking will be used to determine infants' on-line perception of OWs across two different paradigms: wild vs. tame and OW vs. CW. This methodology has been selected due to its capacity for combining visual and auditory stimuli, allowing us to trace infant mapping of sound-meaning correspondences in real-time. While providing an insight into infants' processing of these different forms, the analyses will also uncover some of the issues and questions highlighted in the eye-tracking literature, as discussed above.

4. Eye-tracking study I: Phonological ‘wildness’ in early language perception

4.1 Introduction

Rhodes’ (1994) descriptive paradigm of ‘wild’ and ‘tame’ onomatopoeia serves to reflect the degree of extra-phonetic sound effects used in the production of OWs, which can be altered in order to present a more or less ‘realistic’ approximation of a particular sound. This spectrum of ‘wildness’ calls into question the nature of iconicity, challenging the extent to which its presence in onomatopoeia is dependent on the speaker’s production of a particular word form. Rhodes (1994) defines ‘tame’ forms as being produced within the phonetic norms of the ambient language, adhering to normalised phonological structures that are familiar to the speaker; it can be assumed that these forms do not stand out as any different from the rest of the input. In this sense, tameness represents the maximal conventionalisation of an OW as a lexical form, such as the realisation of *choo choo* as [tʃuːtʃuː], which may originally derive from the sound that some trains make, but which has merely become a referent used to arbitrarily exemplify ‘the sound of a train’. ‘Wild’ forms, on the other hand, make use of the vocal tract’s full capacity in order to appropriately approximate the sound that the speaker is imitating through onomatopoeia. Wild forms draw upon prosody and vocal gestures that are not ordinarily used in the adult language, and it could be posited that these ‘special effects’ render onomatopoeic forms more realistic – or iconic – in the early input. In his brief discussion of sound symbolism in onomatopoeia, Rhodes refers to the production of wild OWs as reflecting a direct mapping between a sound and its meaning, yet he alludes to tame OWs as being sound symbolic “except for one factor, sound symbolism” (p.279). Here Rhodes appears to be suggesting that the sound-meaning relationship in a tame OW is arbitrary, despite the widely-assumed connection between the form and its referent: he states that the form is “simply approximated by an acoustically close phoneme or phonemic combination” (p.279).

However, it could also be suggested that wild forms are more salient, and perhaps more acoustically interesting, than their tame counterparts, owing to the use of prosodic and extra-phonetic features which make them stand out in the input. Rhodes (1994) highlights this in an example of the approximated sound of a sheep, *baa*, which is realised in the tame form as [bɑː], but could be produced with wild features as [bæˤʔæˤʔæˤʔ], with a combination of pharyngeal and laryngeal phonation, as well as impressionistic intonation features reflecting a sheep sound. Similarly, *quack quack* could be realised as [kwæʔkwæk], or again with pharyngeal-laryngeal phonation as [ʔwˤæˤʔ·wˤæˤʔ] in the wild form. These examples demonstrate the contrast between wild and tame forms, and also highlight how wild forms may stand out from the speech stream more clearly through the use of changing phonation and extended segmental duration. Rhodes also discusses wildness in terms of amplitude and pitch, which are mapped iconically onto the form in question in relation to its phonological properties, and which no doubt serve to increase its salience in the input. Here Rhodes (1994) goes beyond onomatopoeia to consider wildness in terms of general sound symbolism (*cheep* and *twitter* are considered alongside *jingle* and *drizzle*, for example), and his approach raises the question of wildness in OW production, and whether or not the wild-tame dichotomy relates to the presence of inherent iconicity in these forms.

Here we must question the idea that it is *inherent* iconicity that makes OWs more learnable in early language development (Imai & Kita, 2014), since wild OWs can be assumed to stand out more clearly than other forms in input speech, including tame OWs. Wildness no doubt increases the perceptual salience of OWs through the use of appropriate sound effects: high pitch and pitch modulations, as well as consonant and vowel lengthening, are known to facilitate word learning (Fernald & Kuhl, 1987; Jusczyk et al., 1992), and so it could be suggested that wildness promotes OW acquisition through a perceptual advantage in the input. On the other hand, we could consider the addition of wild features to be the explicit *rendering* of iconicity through the addition of supra-segmental and prosodic effects, making the form more realistic: even if iconicity is not phonologically present in a form, the use of wildness may bring about iconicity, independent of a word's phonological features. For example, in recordings of mothers reading picture books to their infants (DePaolis et al., 2010[9]), one mother produces the entire phrase *choo choo said the train* with a high pitch and rhythm typical of a 'train sound', and thus both the CW *train* and the OW *choo choo* are produced with wild features. While *choo choo* may be assumed to be iconic regardless of wildness, the

[9] This dataset will be considered in more detail in Chapter 6.

question of whether or not *train* could be considered as iconic in this context is difficult to address. This raises the issue of how we define iconicity, and the extent to which wildness should play a role in the discussion of iconicity in onomatopoeia.

Rhodes' (1994) account of the wild-tame distinction suggests that both sides of this opposition are likely to occur in onomatopoeic production, and so we assume here that infants are exposed to both wild and tame OWs in the early input[10]. Reports from the literature often describe infants' early production of onomatopoeic forms with what could be considered as wild features, appearing to represent the sound that the infant is attempting to imitate: Werner and Kaplan (1963) cite numerous examples from Stern and Stern (1928, cited in Werner & Kaplan, 1963), whose daughter Hilde is reported to have produced 'designatory vocal imitations' (p.101) in reference to 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" (p.101-102). Similarly, Elsen's Annalena (1991) is reported to have produced many forms with such extra-phonetic features, including a 'snuffling' sound to represent a hare, as discussed in Chapter 3. Furthermore, during interactions with a 10-month-old, the current author noted the infant's production of a squealing sound with distinctive pitch modulations in response to his mother's production of *cat*; this was reported by the mother to be a protoword created independently by the infant, owing to the presence of a number of cats in the home. Such instances of infants' depictive imitations raise the question of phonological wildness, as we bear witness to infants' recollection of concepts in relation to the sound that accompanies them.

In this analysis we address the question of whether phonological wildness, as reported in many accounts of infants' early word production, facilitates infants' recognition of OWs in the input, rather than the presence of any inherent iconicity in these forms. Here we draw on Rhodes' (1994) wild-tame paradigm by comparing infants' responses to these two levels of onomatopoeic production. Eye-tracking will be used to determine infants' perception of onomatopoeia in terms of the presence or absence of wildness, in line with the hypothesis that, rather than any inherent iconicity, it is wildness that renders these forms more easily recognisable or identifiable to infants, thus facilitating their acquisition. Rhodes' (1994) 'wild and tame' dichotomy will be applied to the OW stimuli presented, and infants' responses to each condition will be analysed. Here wild

[10] See Chapters 6 and 7 for an analysis of parents' production of OWs in infant-directed speech, including a consideration of wildness.

onomatopoeia are assumed to provide a strong sound-meaning correspondence which may facilitate word comprehension, while tame forms are considered to be less easily identifiable, and thus less easily matched to their target referent. If this is the case, wildness should facilitate recognition of onomatopoeia even when presented in unfamiliar languages: wild features should not be phonologically specified if they truly represent realistic sounds, and thus should not differ despite differences in the target languages. In contrast, if the presence of wildness is the determining factor in infants' recognition of OWs, tame forms will not be recognisable when presented in unfamiliar languages. This goes against claims made by the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014) and by Werner and Kaplan's (1963) model of increasing differentiation. Both of these accounts suggest an inherent advantage for iconic forms such as onomatopoeia in early word learning, with iconicity as a phonologically specified feature which is thus independent of wildness. These claims will be tested further in Experiment II (Chapter 5).

In Experiment I we hypothesise that:

1) Infants will be better able to recognise onomatopoeia when they are presented with wild than with tame features.

2) The addition of wild features will render onomatopoeia in unfamiliar languages more easily recognisable.

3) Tame forms presented in unfamiliar languages will not be easily recognised by the infants.

This study will explore the wild-tame paradigm with the general hypothesis that supra-segmental cues presented in the wild forms facilitate recognition in early language development. When these cues are removed, leaving only the tame forms, onomatopoeia are rendered less recognisable. This would suggest that it is the wildness, and not any iconicity manifested in the phonology of OWs, that leads to infants' early understanding of these forms.

4.2 Methodology

4.2.1 Participants

Nineteen Swedish infants (9 female) between the ages of 14 and 16 months were tested (mean age 461.5 days). A further five infants participated in the experiment but were excluded from the analysis due to fussiness during the eye-tracking procedure (4) or experimenter error (1). Infants were all full-term, and acquiring Swedish as their first language. Nine of the infants were being raised bilingually, acquiring Spanish (1), English (3), Persian (1), Bosnian (1), French (1), Dutch (1) and Norwegian (1) as a second language; none of the infants were reported to have had any exposure to the three unfamiliar languages used in the experiment (Arabic, Mandarin Chinese and Urdu). All caregivers were provided with information sheets about the experiment and asked to sign consent forms (see Appendix I.A and B for relevant documentation). All documentation was provided in Swedish and a Swedish-speaking assistant was present throughout the procedure.

4.2.2 Stimuli and materials

Audio stimuli

The audio stimuli were presented at a volume of approximately 70dB through Creative Inspire T5400 speakers at either side of the screen. Six onomatopoeic words (OWs) – all animal sounds (OW equivalents of COW, SHEEP, DOG, CAT, DUCK and ROOSTER) – were selected for use in the experiment. The chosen stimuli (see Appendix II.A) all appeared on English, Swedish, German and French adaptations of the MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 1994), and so participants were likely to have had prior experience of these words. OWs were recorded in Swedish (the familiar language, L_F), and in three languages that were unfamiliar to the infants (L_U) – Mandarin Chinese, Arabic and Urdu.

Audio stimuli were recorded by native speakers of the four languages, all female postgraduate students in the Linguistics departments of York or Stockholm universities. Each speaker was first asked to produce the OW as they would produce it when speaking to a toddler, adding sound effects as if imitating the animal in question ('wild' W). They were then asked to produce the words with no added prosodic features, keeping to the natural phonology and stress pattern of their native language ('tame' T). The speakers were asked to maintain the conventional full form of the word according to the specific language; words which would normally undergo reduplication were reduplicated (e.g. *quack quack*), while the remaining words were recorded without reduplication (e.g. *cock-a-doodle-doo*).

It was assumed that W forms would be prosodically more salient than T forms due to the nature of this contrast by definition, and so measures were taken to assure that, as far as possible, infants were tested on the effect of wildness, rather than on the prosodic salience of the wild forms. Paired-samples t-tests across W and T stimuli for each of the four languages revealed no significant difference between conditions (W vs. T) in terms of mean pitch range of the six OWs (Arabic: $p=.658$, Chinese: $p=.24$, Swedish: $p=.59$, Urdu: $p=.795$) or mean duration (Arabic: $p=.227$, Chinese: $p=.335$, Swedish: $p=.114$, Urdu: $p=.616$); however, mean pitch was found to be significantly higher amongst the W forms in Urdu ($p=.01$), though not in the other three languages (Arabic: $p=.128$, Chinese: $p=.619$, Swedish: $p=.198$). Any bias towards the Urdu forms will thus be considered in the analysis.

Visual stimuli

Two different photographic images of each of the corresponding animals were selected: the animals were all facing in the same direction, looking towards the infant from the right-hand side. The images were approximately 400x400 pixels in size, presented in colour on a 1280x1024 pixel grey background on either side of the screen in each trial. These were vertically centralised, and set at a distance of 120 pixels from the edge of the screen on either side.

A third image – a cartoon drawing of a crawling baby – was used as a 'centralising' image in the experiment. This was slightly larger than the test stimuli in order to attract the participants' attention, but was also shown looking towards the infant from the right-hand side. This image was presented between the two animal images after the salience phase in order to draw the infants' eyes to the centre of the screen before the test phase (see below).

Similarity

A phonological similarity analysis of the three T_U variants compared with the corresponding T_F variant was carried out in line with Mueller and colleagues' (2003) PSIMETRICA model for measuring phonological dissimilarity. This showed a range of differing similarity scores across the T_U stimuli (see Appendix II.B). The PSIMETRICA model was altered slightly to meet the requirements of infant phonology: phonetic transcriptions of each of the OWs in the three L_U languages were compared against their L_F counterpart on a phoneme-by-phoneme and then syllable-by-syllable basis to ensure that both syllabic and phonological structures were taken into account. In cases where the number of syllables or phonemes did not match across the words, the missing segment was marked as null (Ø).

ROOSTER presented a syllable mismatch in both the Mandarin Chinese (/wə.wə/) and the Urdu forms (/kukəɽũkəɽũ/) when compared with Swedish (/kukəliku/). As both word onsets (Goodman & Jusczyk, 2000) and codas (Echols & Newport, 1992) have been identified as perceptually more salient in infants' early inputs, both of these positions were taken into account in the allocation of scores. Word-initial and word-final syllables were matched for comparison across the forms in order to ensure that the salient phonological features of the words were compared with one another: the penultimate syllable in the longer word was scored as Ø in order to account for this prosodic mismatch. This ensured that salient segments were compared directly, while also taking into account the segmental differences. A feature matrix was used to calculate similarity between phonemes in each word, accounting for a broad range of phonetic possibilities within each feature including nasalisation, lengthening, place and manner of articulation and voicing. Similarity values for each pair of phonemes were calculated as fractions of 14 (total number of possible features): if two phonemes had five of the 14 features in common, they were considered to be 5/14 similar. Mean phonemic similarity was calculated for each syllable, and then an overall mean similarity score was calculated from the average phonemic similarity across syllables. Pairs of identical phonemes were scored as 1, and pairs containing a Ø phoneme were scored as 0. This meant the range of scores appeared as decimals, with 0 being completely dissimilar and 1 being completely similar. Scores ranged from 0.24 to 0.96, with a mean similarity of 0.73. Examples of the similarity scoring procedure can be seen in Appendix II.B.

Four adults, none of them speakers of any of the L_U languages, were tested on their recognition of the L_U stimuli, both W and T, prior to the experiment. Only one of the stimuli was found to be unrecognisable by any of the adults across both W and T forms and was removed from the analysis – Mandarin Chinese *cock-a-doodle-doo* /wə.wə/, which was the form with the lowest similarity score (0.24). Of the remaining six stimuli that were judged incorrectly by at least one of the adults, all were produced in a T manner: the adults were able to identify all W_U stimuli. These results confirmed the suitability of the stimuli used in the infant experiment, as well as supporting the hypothesis that wildness facilitates word recognition.

4.2.4 Procedure

Infants were shown two images – a target image and a distractor – with the accompanying audio stimuli differing by wildness (wild vs. tame) and familiarity (familiar vs. unfamiliar). The experiment was controlled using E-Prime (Psychology Software Tools, 2012), with the visual stimuli presented using a 17" Tobii Studio 1750 eye-tracking monitor in a sound-attenuated room. Caregivers held their infant on their lap in a chair placed 60cm away from the screen, and a five-point infant calibration was taken for each participant before the experiment began. The experimental procedure lasted approximately four minutes, during which time the caregiver wore sound-insulating headphones playing music from a Swedish radio station.

The experiment consisted of a salience phase and a test phase: pairs of images were displayed on the screen, and the salience phase – during which time the infants could familiarise themselves with the visual stimuli – lasted for 4000ms, before a centralising image of a baby appeared between the two images which served to 'reset' the infants' eye-gaze prior to the test phase. The centralising image disappeared automatically when the infant fixated on it (or after 4000ms if the infant did not fixate), leaving only the target and distractor images on the screen, and the audio stimuli was played immediately after offset of the centralising image. The test phase lasted for 3000ms. After the experiment infants were rewarded with a certificate and parents were asked to complete a Swedish CDI questionnaire.

Each infant heard a total of 24 OWs: six OWs each presented once in each of the four conditions (W_F, W_U, T_F and T_U), with all three unfamiliar languages distributed equally across the stimuli. The order of data output and the target's location on screen was randomised using E-Prime. Selection of the distractor image was partially randomised in E-prime according to the 'size' of the animal depicted in the target image: to ensure against confusion between the images (e.g. large animals such as sheep and dog, small animals such as duck and rooster), animals were grouped into two categories –'small' and 'large'– and for each trial the distractor image was chosen from the opposite category to avoid ambiguity.

4.3 Results

The infants' responses were analysed for proportion of fixations to target (reflecting understanding of the audio stimuli) and response latency (reflecting speed of recognition). For both of these measures, fixations during test phase were analysed within a window of 300-1800ms after onset of the stimulus, reflecting the expected eye movement latencies for 14-16 month-old infants, adjusted from Swingley and Aslin's (2000) analysis of slightly older infants. The proportion of looking towards the target was calculated for each trial as a percentage of the total fixation time for both target and distractor, and the mean proportion of looking time towards the target image was calculated for each infant in each condition.

The first point to be observed is infants' responses to the L_F and L_U stimuli. Mean fixation proportion across stimuli was calculated for each infant in both L_F and L_U conditions, and binomial tests showed that recognition exceeded 50% in the L_F condition (n= 19, p=.019), but not in the L_U condition (n= 19, p=.36). This indicates that, when averaged across stimuli, infants were generally able to recognise the L_F stimuli but not necessarily the L_U stimuli. However, average looking time in both conditions was close to 50% (L_F: M=.52, SD=.07; L_U: M=.51, SD=.06).

4.3.1 Wildness and fixation time

A Shapiro-Wilk test confirmed normality across responses (D(19) = .966, p=.697), and thus parametric statistical analyses were performed. Two linear models was constructed to test for any effects of age and sex on the infants' overall mean fixation time: no effect was found for sex (F(1, 16) = .71, p=.41) or age (F(1, 16) = .005, p=.94). A linear mixed-effects model was generated with the fixed effects of type of stimuli (CAT, COW, DOG, DUCK, SHEEP and ROOSTER). Subject was included as a random effect with by-subject random slopes for the effect of type of stimuli; by-item random slopes were not included as item was not modelled as a random effect here. P values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. No effect was found for type of stimuli on the infants' responses ($\chi2$ (1) = .248, p=.62). The same model was then used to compare infants' responses across reduplicated and non-reduplicated stimuli, with reduplication and subject as fixed and random effects, respectively. By-subject and by-item random slopes for the effect of reduplication were included in this model. No difference was found between infants' fixations to target across reduplicated and non-reduplicated stimuli ($\chi2$ (1) = 0.017, p=.89). Item and reduplication can thus be discounted as confounding variables in this analysis.

Linear mixed-effects models were then generated to determine whether wildness (W vs. T) or familiarity (L_F vs. L_U) had any effect on the infants' fixation times to the target image. Fixation to target was tested as the dependent variable, with wildness and familiarity as fixed effects and subject and item as random effects. By-subject and by-item random slopes were included for the effect of wildness. P values were generated using likelihood ratio tests, and showed no significant effect for either wildness ($\chi2$ (1) = 0.194, p=.66) or familiarity ($\chi2$ (1) = 0.578, p=.45) on the infants' fixation times. As Figure 4.1 shows, we see hardly any difference in responses across conditions, with slightly longer looking in the L_U condition. Furthermore, the responses in the L_F condition hover around 50% for both W and T stimuli.

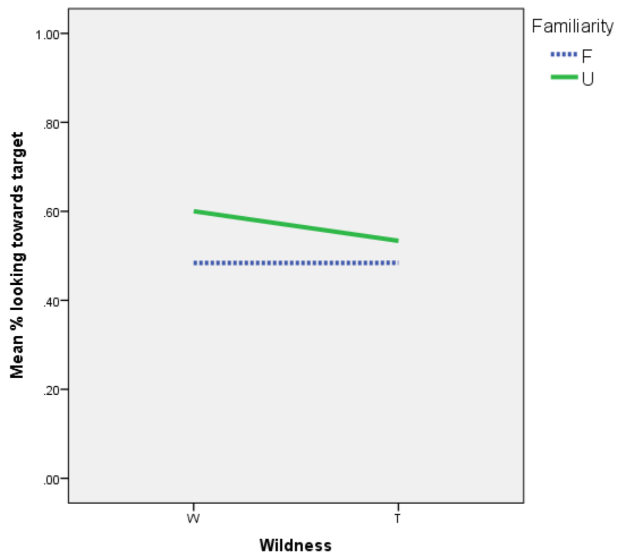

Figure 4.1: Wildness x familiarity.

4.3.2 Wildness and response latency

Results were then analysed to compare the infants' response latency across conditions. Trials in which the infant was fixating upon either the distractor or the centralising image at the onset of the analysis period were considered, accounting for 29% of test trials across the full dataset. In these trials, the amount of time that the infant took to shift from the distractor or centralising image to the target was calculated, reflecting the response latency in seconds. Data from three infants was omitted from the analysis due to a failure to provide a response in each of the four conditions, leaving a sample of 16 infants for analysis.

A Shapiro-Wilk test confirmed normality across responses (D(56) = .962, p=.073). Results were modelled in R using the same two linear mixed-effects models generated above, this time with response latency (seconds) as the dependent variable. Likelihood ratio tests showed wildness to have no effect on infants' response latency ($\chi2$ (1) = 0.585, p=.44), though familiarity had a near-significant effect ($\chi2$ (1) = 3.52, p=.06). Surprisingly, results show that infants looked towards the target image around 250ms faster in the L_U condition, going against our hypothesis that infants would show less recognition in this condition. This is shown in Figure 4.2.

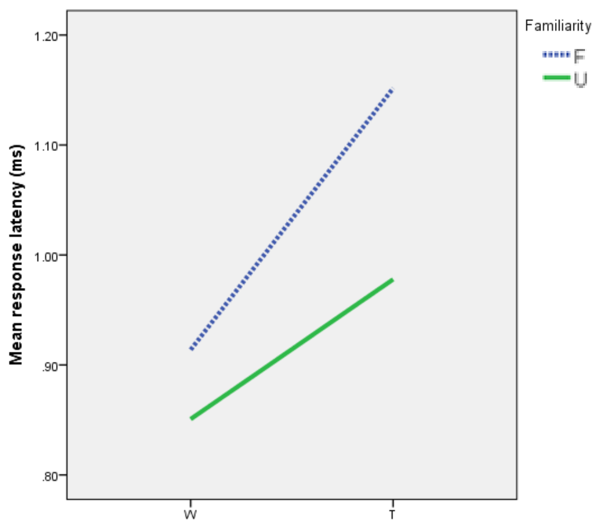

Figure 4.2: Wildness x familiarity.

4.3.3 Language experience

The infants' knowledge of the stimuli was then considered as a factor in our results, in order to assess whether production ability may have determined their responses. Vocabulary sizes (as reported by the caregivers when filling out the CDI questionnaires) ranged from four to 117 words overall, with number of words in the 'sound effects and animal sounds' category varying from zero to ten. No correlation was found between the infants' age in days and overall reported word production (r = .208, n= 19, p=.992), but a strong correlation was identified between the infants' overall vocabulary and the number of OWs that they were reported to produce (r = .649, n= 19, p=.003). For each of the six targets, infants were grouped into one of two categories – 'can produce' and 'cannot produce' – according to their ability to produce each form. Fifteen infants had words in both categories, while four infants were unable to produce any of the stimuli and thus could not be considered in this analysis; none of the nineteen infants were able to produce all of the stimuli used in the experiment. Overall, 15 infants provided sufficient data across all variables and both categories to be considered in this analysis. A linear mixed-effects model compared infants' responses to words that they reportedly could and could not produce. Proportion of fixations to target was measured as the dependent variable, with production ability (can vs. cannot produce) as a fixed effect and participant as a random effect with by-subject random slopes for production ability. No significant effect was found for the infants' production ability ($\chi 2$ (1) = 0.12, p=.78): infants did not fixate longer upon hearing targets that they could produce. Wildness and familiarity were then added to the model as fixed effects, to test whether there was any difference in infants' responses across these two variables. Again, no significant effect was found ($\chi 2$ (1) = 0.23, p=.63). Results for mean response latency across these two categories were available for only two infants, and so no comparison could be made.

4.3.4 Similarity scoring across stimuli

The differing similarity scores across L_U stimuli were then considered in order to investigate the apparent bias in the L_U condition. A Pearson product-moment correlation coefficient was used to compare mean proportion of fixations to the target image with the similarity score of the 18 stimuli presented in the L_U trials. No correlation was found between the similarity of the stimuli to the L_F target and infants' fixations to the target image in either W (r = .325, n= 18, p=.203) or T trials (r = -.263, n= 18, p=.308). Neither was there a correlation between similarity score of the target and mean response latency across all infants for the 12 stimuli for which response latency data were available (W: r = -.452, n= 12, p=.141; T: r = .296, n= 12, p=.350).

Mean similarity scores differed across languages (Arabic: 0.82, Chinese: 0.64, Urdu: 0.73), and so infants' responses in the L_U condition were considered with target language as a factor, thus also accounting for the higher pitch identified in the Urdu stimuli, as discussed above. A linear mixed-effects model was constructed in R with fixation proportion as the dependent variable, language (Arabic vs. Chinese vs. Urdu) and wildness as fixed effects and subject as a random effect, including by-subject random slopes for language. Language had a significant effect on infants' fixation times in the L_U condition ($\chi 2$ (1) = 5.29, p=.021): as shown in Figure 4.3, infants showed the longest fixation times upon hearing the Arabic stimuli. This corresponds to the higher mean similarity score observed in these stimuli, suggesting that infants may have looked longer to target upon hearing OWs that were most familiar in the input. However, this trend does not follow with the Chinese and Urdu stimuli, and as Figure 4.3 shows, responses to the Swedish stimuli do not correspond here either: if similarity were a factor in infants' responses, we would expect to see a gradient difference between fixation times in relation to expected recognition across the L_F stimuli and the three L_U languages. Instead, Swedish has the lowest mean fixation proportions, followed by Urdu and then Chinese and Arabic.

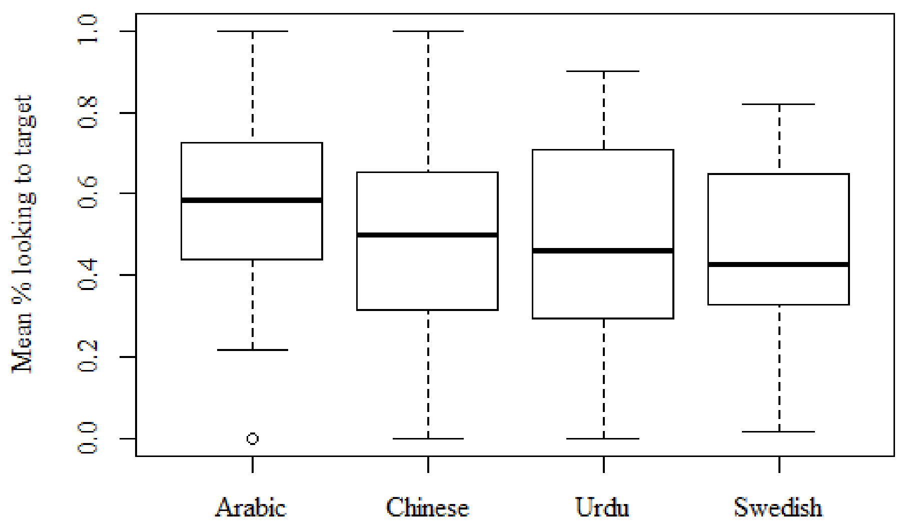

Figure 4.3: Fixations across the four languages.

As the lowest L_U fixation times were observed in the Urdu trials we can confirm that the higher pitch of these stimuli did not cause any bias towards these stimuli in particular. The model showed no effect for wildness (p=.68). The same model was then generated with response latency as the dependent variable[11], and no effect was found for language (p=.19) or wildness (p=.73).

4.3.5 Prosodic effects of L_U stimuli

Preference for the L_U condition has been observed across infants' responses here, and this differs across the three language conditions. Similarity to the L_F stimuli could be posited as a factor, but inconsistencies across the four languages suggest that this may not be the case. Prosodic differences between W and T forms were controlled for in the initial analysis, but no measures were taken to control for the effects of prosody across

[11] Random slopes could not be included here due to the small amount of response latency data.

languages: here we raise the question as to whether the individual L_U forms could be more salient than the L_F forms regardless of any wild/tame distinction. Linear mixed-effects models in R tested mean f0, pitch range and duration of the stimuli as dependent variables, with language (Arabic vs. Chinese vs. Urdu vs. Swedish) and wildness as fixed effects and item as a random effect. By-item random slopes for language were also included.

Mean pitch

Language was found to have a strong effect on the mean pitch of the stimuli ($\chi2$ (3) = 13.42, p=.004): the L_F stimuli were on average around 73Hz lower in pitch than the L_U stimuli. As shown in Figure 4.4, Urdu had the highest mean pitch, closely followed by Arabic: this may provide some suggestions as to why infants fixated longer to the target in the L_U condition, and as shown in Figure 4.3, this was most prominent for the Arabic stimuli. No effect was found for wildness here (p=.79).

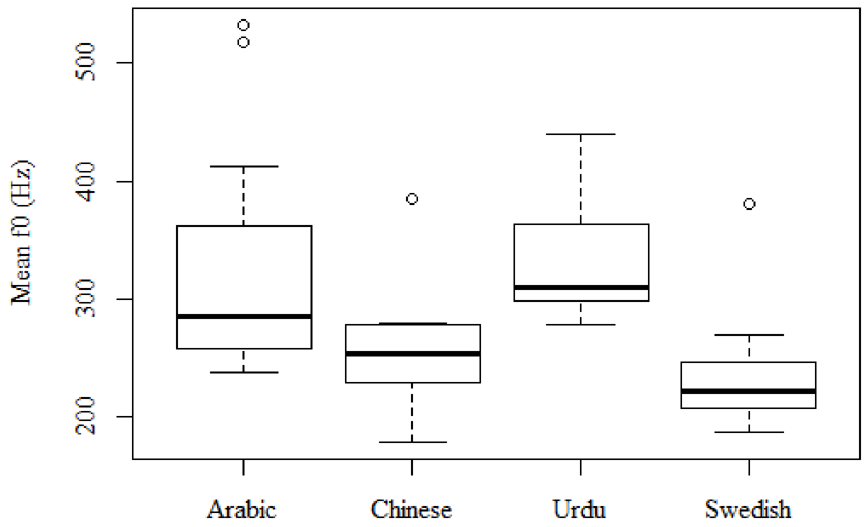

Figure 4.4: Mean f0 of stimuli across the four languages.

Pitch of each of the six stimuli was then considered in relation to fixation proportion. A linear model comparing infants' fixation proportion as a function of pitch of each of the individual stimuli revealed a significant effect for pitch on the infants' fixations (F(1, 33) = 6.39, p=.016): infants fixated longer on the target image upon hearing stimuli with a higher f0. This is shown in Figure 4.5, where we also observe a significant correlation between pitch of the individual stimuli and the mean fixation proportion to those stimuli (r = .403, n=35, p=.016).

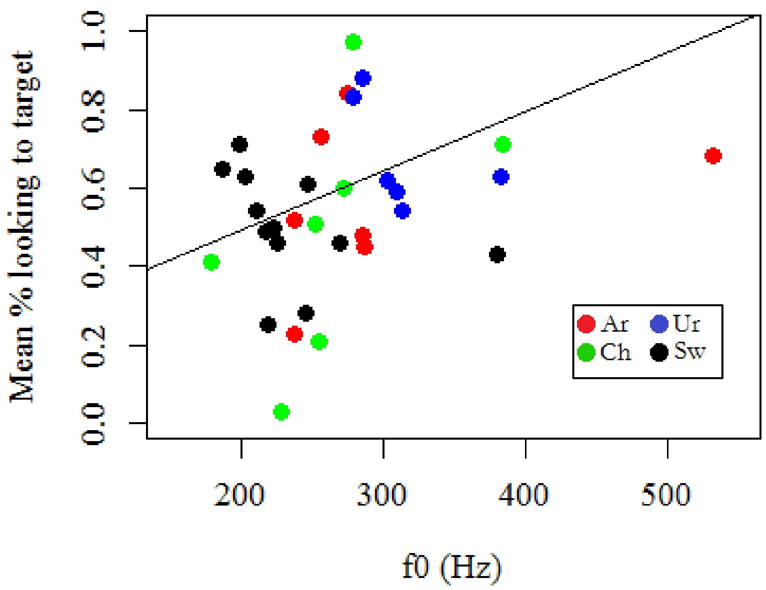

Figure 4.5: Pitch stimuli against mean fixation proportion across infants.

Duration

Language was also found to have a significant effect on duration of the stimuli ($\chi 2$ (3) = 11.296, p=.01), with L_F stimuli on average 858ms shorter than L_U stimuli. This difference is clearly shown in Figure 4.6. This suggests that stimuli duration may also play a role in infants' response times. Again, no effect was found for wildness (p=.08).

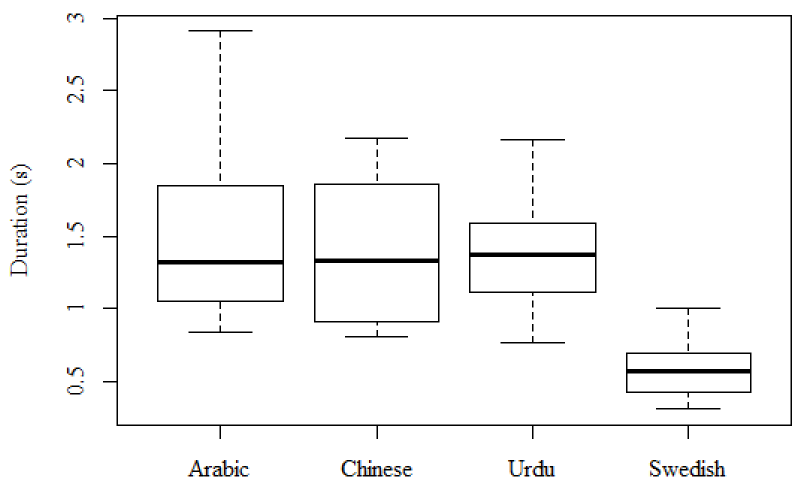

Figure 4.6: Mean duration of stimuli across the four languages.

A linear model analysing infants' fixations to target as a factor of target duration showed no effect for the duration of the stimuli (F(1, 33) = .34, p=.57). We can see from Figure

4.7 that stimuli of around 1.5s in duration generated the longest looking times in all, and that fixation proportion to target decreases after this point. Here we may be observing a ceiling effect in infants' responses, whereby extended word duration affects infants' looking time up to a certain point.

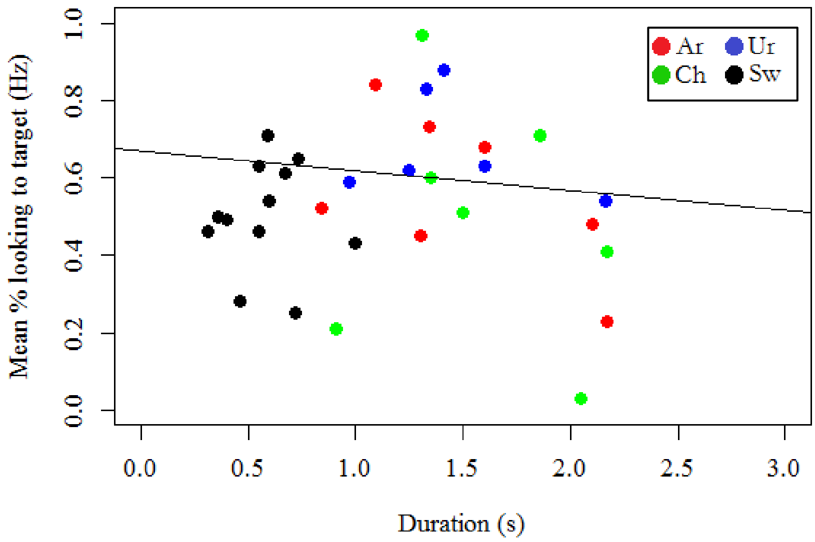

Figure 4.7: Duration of stimuli against mean fixation proportion across infants.

Finally, language had no effect on the pitch range of the stimuli ($\chi2$ (3) = 4.47, p=.21), and nor was there any effect for wildness here (p=.48).

4.4 Discussion

Neither wildness nor familiarity played a role in infants' responses during this experiment; contrary to our original hypotheses, wildness did not have any advantage over tameness, and did not facilitate recognition of unfamiliar forms. Furthermore, the stimuli presented in unfamiliar languages consistently generated the longest and fastest responses (shown in Figures 4.1 and 4.2, respectively), though we did not see any statistical significance in this regard either. In all, none of the three hypotheses were supported by these results, and the fact that we did not see any effect of recognition in the L_F condition suggests that responses may have been driven by factors other than wildness or familiarity. Initial analyses showed that infants' mean looking proportions were around chance (50%) across both L_F and L_U conditions, highlighting some problematic features of the data. The presentation of four different voices across the stimuli, as well as large prosodic differences both within and between speakers, no doubt led the results to converge to the extent that our research questions were not answerable. However, while a number of confounding issues have obscured our results, still some interesting points can be raised following this analysis.

The faster responses in the L_U condition are surprising and contradict all proposals drawn in the hypotheses. Indeed, results from the literature suggest that we should expect *slower* response latencies for the L_U forms, as shown by Swingley and Aslin (2000) in their study of infants' responses to correctly-pronounced and mispronounced words. Infants of 18-23 months showed significantly slower responses to mispronounced words, including changes to the voicing of the onset consonant (*dog* vs. *tog*) or the medial vowel (*car* vs. *cur*), or changes in place and manner of articulation (*baby* vs. *vaby* and *ball* vs. *gall*). The variable similarity across L_F and L_U stimuli mirrors the changes used in Swingley and Aslin's analysis, with contrasts in onset voicing (Swedish /kvakvak/ vs. Chinese /gɑgɑ/ DUCK), nasalisation (Swedish /muː/ vs. Urdu /bʌː/ COW) and vowel changes (Swedish /muː/ vs. Arabic /mɵː/ COW), as well as more complex differences (Swedish /vɔvɔv/ vs. Arabic /hawhaw/ and Chinese /wʌŋwʌŋ/ DOG). However, an important difference between Swingley and Aslin's (2000) methodology and the present study could account for this unexpected difference in our results, as half of the stimuli in our experiment were presented with wild features. This may have made recognition easier for the infants owing to the extra linguistic information that was available for each word, as wild features were similar across the stimuli despite phonological differences across languages (i.e. similar wild features were used in the production of DOG even though the phonological form differed across the four languages).

Discrepancies in the salience of the L_U forms might also account for these trends: a prosodic advantage was found across these stimuli, which may explain why we see a looking bias in the L_U condition. The L_U forms were significantly higher in pitch and longer in duration, which may have caused them to draw infants' attention to the target image more successfully during the experiment; both of these features have been shown to be advantageous to infants' perception of language in IDS (Fernald & Kuhl, 1987; Juscyzk et al., 1992). Figure 4.5 confirms that this may be the case, as we see that those stimuli with the highest pitch generated the longest looking times. The same cannot be said for word duration (Figure 4.7), though we propose that a ceiling effect may be distorting the results here. Furthermore, of the six stimuli that have a duration of above two seconds, three are COW and two are ROOSTER. Infants might also be biased against these stimuli since they may be less likely to recognise or show interest in these animals than the potentially more familiar DOG, CAT or DUCK stimuli.

The use of reduplication also differs substantially across L_U and L_F stimuli: only two of the six L_F forms are reduplicated, while all of the L_U forms contain reduplication. No effect was found for reduplication on infant fixations, but it is likely that the presence of reduplication in the L_U forms contributes to their longer duration, and may also lead to the (non-significant) looking bias that we see in this condition. By its very nature reduplication presents the repetition of a segment within a word, which may contribute to the faster recognition of L_U forms observed in Figure 4.2, as the input is doubled within a single token. Indeed, the presence of within-word repetitions is thought to be advantageous in language processing: Gervain and colleagues (2008) found that neonates were able to distinguish between words which contained repetitions (AAB words, such as *mubaba*) and those that did not (ABC words, as in *mubage*), but the results did not hold when those repetitions were not in direct sequence with one another (i.e. when an ABA word such as *bamuba* was contrasted with an ABC word). The authors suggest that there may be an implicit advantage for such repetitions in early language processing, referred to by Gervain and colleagues as a "perceptual repetition detector" (p.14226). These findings are also consistent across adults, as shown by Endress and colleagues (2007) in a number of artificial grammar learning tasks. Based on these analyses, it is suggested that repetition (and thus also reduplication) is perceptually salient to infants in the early input, and may facilitate lexical acquisition – hence the common presence of reduplication in the early output (Gervain et al., 2008; Gervain & Werker, 2008). In this respect, the combination of higher pitch, extended duration and reduplication in the L_U stimuli no doubt serves to provide a processing advantage for these forms in the early input. In Chapter 2 we observed how reduplication is a common feature of OWs, and its potential for maintaining infants' attention, as well as providing a 'double input' of the same form, should not be overlooked.

The infants' bias towards the L_U stimuli paired with their surprisingly low interest in the L_F stimuli could also be explained by a familiarity effect. In the L_U condition, infants were presented with a slightly distorted version of a familiar word form, which may have prompted longer looking times (Fantz, 1964). Normally cited in experiments involving the head-turn preference procedure (HPP), this effect is thought to relate to the influence of memory representation on infants' responses (Roder et al., 2000). When representations are incomplete in the early phases of processing, a preference for the familiar can be found in an infant's responses, which persists if *complexity* is added to the stimulus. That is, those stimuli with enough complexity to be deemed interesting, yet not so much so that they cannot be understood, will continue to elicit a familiarity response (Kidd et al., 2012). Indeed, this could explain the infants' responses to the L_U stimuli, which may have been sufficiently familiar and complex to generate longer fixations, owing to both the high similarity and the "intermediate surprisingness" (Kidd et al., 2012) found in the phonology of these forms. Similar results were observed in Vihman and colleagues' (2004) HPP experiment, where 11-month olds showed longer listening times for mispronounced familiar words than unknown words, and Swingley and Aslin (2002) found the same response among 14-month olds. Here, L_U stimuli present an alternation of the established form, and may elicit a response to the "intermediate surprisingness" of a word form which in most cases is slightly different from the infants' phonological representation of the word. Furthermore, 50% of the 18 L_U forms used in this experiment have an onset consonant that differs from that of the L_F form[12] (see Appendix II.A); onset consonants have been shown to play an important role in infant word perception (Goodman and Jusczyk, 2000), and may have caused the infants to question their representations of OWs, prompting a longer looking time (i.e. a familiarity response) towards the target in this condition.

When the infants' responses are compared with results from the literature, response latency measures in the T_F condition correspond to those observed by Fernald and her colleagues (1998) with regard to infants of 14-16 months: a similar experimental design showed a mean response time of 995ms for 15-month-olds in matching audio stimuli to a target image. This is only 6ms slower than the mean responses observed in the T_F condition here, while responses to the W stimuli are on average 207ms faster than in Fernald and colleagues' results. Furthermore, as responses to tame forms did not differ substantially from the response latencies to non-OW forms observed in Fernald and colleagues' (1998) analysis, we could suggest that this might reflect a prosodic advantage for wild OWs – this would contradict the blanket assumption that all OWs are iconic,

[12] Five of these nine onset consonants differed only in voicing or nasalisation, while two differed in manner of articulation and two didn't correspond at all.

regardless of the use of wild features. However, the lack of non-OW stimuli in the present study (as well as the lack of any effect in our data) leaves this comparison open to debate, since empirical claims cannot be made through the joint consideration of these two separate experiments.

Tentatively, two confounding conclusions can be made regarding these results. First, we could interpret the lack of difference between W and T stimuli to support a role for intrinsic iconicity in OWs. While the comparison with Fernald et al.'s (1998) analysis drawn above shows that infants' responses to tame OWs do not differ from responses to non-onomatopoeic words, the lack of advantage for wildness in the present study may suggest that iconicity is inherent in onomatopoeia regardless of wild features. The comparison of these two experiments does not suffice in providing empirical evidence either in support of or contradictory to Imai and Kita's (2014) sound symbolism bootstrapping hypothesis, and so a follow-up experiment in Chapter 5 will allow us to address this possibility more clearly. Second, we could suggest that the salient prosodic features of the L_U forms are a main factor, which draw infants' attention to OWs in the input when wild features are used in OW production. Our initial analyses showed no prosodic advantage for W forms over their T counterparts, but overall we do see a clear prosodic advantage for those stimuli – regardless of language or wildness – that are presented with the most prosodic salience. This is in line with previous research demonstrating the importance of salient features in IDS (see Cristia, 2013, for a comprehensive review), and will be considered in terms of OW production in Chapter 6.

It is not possible to truly separate the consideration of iconicity from that of prosody in these results, and in order to escape the circularity of this question it is necessary to control for the prosodic characteristics of the stimuli across the board. The inclusion of wildness makes it difficult to avoid a prosodic bias towards the wild forms, and it is unclear whether this would be possible to control for in any future experiments. Furthermore, the comparison of four different languages using four different speakers makes it difficult to thoroughly interrogate the question of iconicity independently of prosody: it would be necessary to separate the questions of wildness and familiarity across two individual experiments if these issues were to be overcome.

4.5 Conclusion

While familiarity and wildness had no effect on infants' responses, the results showed an interesting trend in terms of prosodic salience, which was clearly driving infants' responses during the procedure. Our results show a clear advantage for mean pitch and possibly also duration in infants' responses to the stimuli: these findings may be central to this thesis as a whole, as this leads us to further consider the nature of OW production in the infant input. This will be analysed further in Chapter 6.

Overall, these results leave a number of unanswered questions, since we were unable to separate the question of iconicity and the potential role for wildness from the influence of prosody on early infant perception. It remains unclear as to whether wildness facilitates infants' recognition of an OW, and whether the tame form has an iconic advantage independent of any wild features. Furthermore, as this experiment only considered infants' responses to OWs, we are not able to draw any conclusions regarding whether or not OWs possess any intrinsic iconic advantage, over and above the presence of wildness. This will be analysed in Chapter 5, where responses to OWs will be compared with CWs.

5. Eye-tracking study II:
Inherent iconicity in onomatopoeia?

5.1 Introduction

In Experiment I we considered the role of wildness in infants' perception of onomatopoeia to identify whether this had a facilitative effect on OW recognition. Infants showed no difference in their responses to wild and tame stimuli, leading us to conclude that wildness does not play a role in infants' perception of OWs. Discrepancies in the analysis led to some mixed findings and thus the question of wildness and iconicity could not be discussed with sufficient rigour. Indeed, the potential perceptual advantage for OWs owing to the presence of iconicity in these forms has not yet been fully addressed, and as concluded in the previous chapter, it is clear that a consideration of wildness alone cannot account for this.

Iconicity has been shown to be inherent in Japanese mimetics: even non-Japanese speakers – infants as well as adults – have been found to correctly match sounds to meanings when presented with unfamiliar mimetic words (Imai et al., 2008; Iwasaki et al., 2007). This is discussed widely in the iconicity literature, where the presence of segmentally-specified sound symbolic features (Maurer et al., 2006; Monaghan et al., 2011) is thought to facilitate the recognition of iconic forms. However, while iconicity is generally assumed to also be inherent in onomatopoeia (Monaghan et al., 2011; Nygaard et al., 2009), no discussion has taken place to consider how meaning may be phonologically present in these words. As already discussed, Rhodes (1994) opposes this view by suggesting that sound symbolism in onomatopoeia is present only in wild forms, but this is not considered in any other discussion of onomatopoeic words. Without taking wildness into consideration, many researchers appear to be suggesting that iconicity – and thus the iconic advantage – is manifested in the phonology of OWs.

In Experiment II we attempt to address these assumptions explicitly, through a comparison of infants' responses to onomatopoeic and non-onomatopoeic forms. The possibility of a prosodic advantage will be controlled for through the presentation of exclusively tame OWs, in order to separate the question of iconicity from that of wildness and prosodic salience. Infants' looking time and response latencies will be measured in order to determine whether OWs enable more efficient recognition and mapping in early language development than their equivalent CWs: longer fixations on the target image upon hearing the OW forms over the CWs would suggest that the iconic properties of OWs may facilitate infants' recognition of these forms in early word learning. In addition, shorter response latencies would demonstrate a processing advantage for OWs over their CW equivalents. Eye-tracking will be used to determine infants' fixation time and response latency to a target image in relation to OW and CW stimuli.

5.2 Methodology

5.2.1 Participants

Parents were recruited through an advert in a local magazine, through social media and by word of mouth. Forty 10- and 11-month-old infants took part in the experiment overall, of which 13 were excluded from the analysis due to equipment failure (n= 5), infant fussiness (n= 3) and calibration problems (n= 4), leaving a total of 27 infants for analysis (14 females, mean age 328 days). All of the infants were acquiring British English, and one of the infants received some exposure to Mandarin alongside English. No developmental difficulties were reported, and all but two infants were reported to have had full-term gestational periods: the two infants in question were dizygotic twins, whose results did not differ from those of the other 25 infants in the study. On average the infants looked at the eye-tracking monitor during 57% of the trials. All caregivers were provided with an information sheet explaining the procedure and were asked to sign a consent form (see Appendix I.C and D for relevant documentation).

5.2.2 Stimuli and materials

Audio stimuli

Six OWs from the 'sound effects and animal sounds' section of the Oxford Communicative Development Inventory (CDI, Hamilton et al., 2000) were selected for use in the test trials, each with a CW counterpart from elsewhere on the CDI. These words were chosen on the basis that they would be familiar to 10 and 11-month-old infants. Two further OW-CW pairs from the CDI were selected to use as filler trials, to provide different audio and visual stimuli in order to help maintain infants' interest during the experiment. Stimuli are detailed in Table 5.1, below. Audio stimuli were recorded by a female speaker of Northern British English, to reflect the dialect spoken in York, UK, where the experiment was carried out. The speaker – a linguist from the department at York – was asked to produce each word in the carrier phrase "Where's the [target]?", and it was specified that OWs and CWs should be produced with the same pitch contour. Word duration ranged from .524ms to .824ms across both OW and CW stimuli.

Table 5.1: OW and CW stimuli used in the experiment.

OW	CW
Baa	Sheep
Meow	Kitty
Moo	Cow
Vroom	Car
Quack quack	Ducky
Woof woof	Doggie
Cock-a-doodle-doo	*Cockrell*
Choo choo	*Train*

Filler stimuli are marked in italics.

Multiple tokens of each target were recorded, and the final stimuli were selected based on the closest match for each OW-CW pair in terms of pitch and duration, in order to ensure that neither member of each stimulus pair stood out as more salient than its counterpart. Mean pitch, duration and pitch range of the stimuli were modelled using a linear mixed-effects analysis in R, in order to account for any differences in prosodic salience across stimuli. Type (OW vs. CW) was included as a fixed effect and target (SHEEP vs. CAT vs. COW vs. CAR vs. DOG vs. DUCK) as a random effect; random slopes were not possible for this small set of data. No effect was found for either factor on mean pitch ($p=.99$) or duration ($p=.269$), though type did have an effect on pitch range ($\chi2(1) = 4.2$, $p=.04$): OW stimuli had a wider pitch range than CW stimuli by an average of 29Hz. This discrepancy is not considered to be substantial, but will be accounted for in the analyses that follow. Paired-samples t-tests confirmed that this discrepancy in pitch range did not have an effect across individual OW-CW pairs ($t(5) = 1.868$, $p=.121$).

Each test pair was also matched for number of syllables; this was not possible in the filler pairs but was not considered to be an issue. For three of the stimuli – *kitty*, *doggie* and *ducky* – diminutives were selected for the CW in each pair, as this was considered to be a common production of these target words. This was matched to the syllabification of the corresponding OWs, with the inclusion of reduplication in *quack quack* and *woof woof*.

Visual stimuli

Two different photographic images were selected to match each OW-CW target, and one image each for the filler pairs. It was ensured that the six animals all stood facing towards the right-hand side of the screen with their head turned towards the infant, while the car and train images were presented with the front-end of the vehicle towards the right-hand side of the screen. Two images of approximately 400x400 pixels in size were presented in colour, side-by-side, on a 1280x1024 pixel grey background. These were vertically centralised, and set at a distance of 120 pixels from the edge of the screen on either side.

Apparatus

The experiment was controlled using PsychoPy Experiment Builder (Peirce, 2007) and was run through Tobii Studio, presented on a 17" Tobii Studio T60 eye-tracking monitor sampling at 60Hz. The experiment was set up in a darkened booth with speakers installed in the walls of the booth at each side of the monitor, and stimuli were played at a volume of approximately 70dB. A video camera positioned above the eye-tracker allowed the experimenter to view the procedure from the adjoining room, providing information regarding the infants' comfort as well as orientation towards the screen.

5.2.3 Procedure

An image from the children's TV programme *Teletubbies* was displayed on the eye-tracking screen prior to the infant entering the experimental booth. This distracted the infant while the experimenter set up the procedure and drew their attention towards the screen ready for the start of the experiment. Caregivers wore foam earplugs and headphones playing multi-talker babble, as well as visibility-blocking glasses. These ensured that the caregivers did not influence their infant's responses during the experiment. Infants were held on their laps in a chair placed 60cm from the screen.

A five-point infant calibration was taken, after which point the experiment began. This lasted approximately 3.5 minutes in total and consisted of 24 randomised test trials and four filler trials: a filler trial was presented at the start of each phase, followed by six test trials. This was repeated for four phases of the experiment. Each infant heard 14 OWs and 14 CWs, presented in a random order which was counterbalanced so that the target image was displayed equally on both the left- and right-hand sides.

Infants were first presented with the two images on screen in a 3700ms familiarisation phase, before the audio stimuli ("Where's the [target]?") was played. The images remained on screen until 3500ms after the onset of the audio stimuli. After the experiment infants were rewarded with a specially-designed Babylab t-shirt, and parents were asked to complete the Oxford CDI questionnaire (Hamilton et al., 2000). Caregivers were specifically asked to indicate whether or not they used OWs in the home, and this was marked on the CDI form accordingly.

5.3 Results

Data was analysed from a window of 350 to approximately 2500ms after onset of the audio stimuli. This analysis period was selected based on Swingley and colleagues' approach (Swingley, Pinto & Fernald, 1999; Swingley & Aslin, 2000), who typically adjust the assumed latency period between target onset and response according to the age of the infant. Twelve-month olds have been found to show mean saccade latencies of around 290ms (Canfield et al, 1997), and so this was extended here to allow a slightly longer response time for 10-11 month old infants. The offset of the analysis window was also adjusted from that of Swingley and Aslin's (2000) study in order to allow for slower and longer responses from the younger infants.

As in Experiment I, proportion of looking towards the target image was calculated for each trial as a percentage of the total fixation time for both target and distractor, and a mean looking time was calculated for each infant for both OW and CW stimuli across the six targets. Five infants were excluded from the dataset owing to a lack of results across conditions, leaving 22 infants for analysis. In order to assess recognition of the stimuli, positive difference scores were calculated across the results, in line with Bergelson and Swingley's (2012) analysis of six to nine-month old infants. Here, differences in fixation proportions were used as an indication of whether the infants recognised the six stimuli and their corresponding OW and CW labels: this was calculated as the difference in fixation proportion to image A when A was the target in relation to the fixation proportion to image A when A was the distractor. For example, infants' proportion of looking to the CAT image upon hearing *kitty* or *meow* relative to their proportion of looking to the CAT when this was the distractor image. Bergelson and Swingley (2012) state that positive difference scores reflect word understanding while also taking into account the infants' preference for any of the visual stimuli. Eighteen of the 22 infants showed a positive mean difference score across the results (M=.045, SD=.15), and binomial tests confirm this to be significantly above the chance value of zero (p=.004). The infants showed positive performance on all six of the individual stimuli, indicating understanding across all word-image pairs and confirming that the infants did not show any looking preference during the test trials.

As in Experiment I, trials in which the infant was fixating on the distractor image at the onset of the test phase were also analysed, in order to determine response latencies from distractor to the target image across the two conditions. Sixty three percent of all trials were included in the response latency analysis.

A Shapiro-Wilk test confirmed normality for both OW (D(22) = .975, p=.827) and CW (D(22) = .952, p=.346) responses, and so parametric tests were used throughout the analysis. To test for any effect of age, sex or language ability, a linear mixed-effects model was generated in R with mean fixation proportion as the dependent variable, subject and item (stimuli) as random effects and by-subject and by-item random slopes for the effect of language ability. This confirmed that the infants' age (p=.37), sex (p=.23) and language ability (number of words reported in the CDI, p=.74) had no effect on their overall mean fixation proportion when these were considered as fixed effects.

5.3.1 Fixation proportion

A linear mixed-effects model was then generated with fixation to the target image as the dependent variable, and the fixed effects of condition (OW vs. CW). Subject and item were included as random effects, with by-subject and by-item random slopes for the effect of condition. P values were generated through likelihood ratio tests of the full model with the effect in question against the model without the effect in question. Infants were shown to fixate 8% longer in the OW condition, but this difference was not significant ($\chi2$ (1) = 3.28, p=.066). Furthermore, average proportion of fixations to target remained close to chance (0.5) throughout, and looking time was on average only 75 milliseconds longer in the OW condition.

Responses in the OW condition were found to be consistently higher across all targets but SHEEP (see Figure 5.1). However, when the stimuli were considered individually, OWs DOG (M=.64), CAT (M=.56) and DUCK (M=.54) were the only stimuli to generate responses which were clearly above chance (50%, see Figure 5.1), corresponding to reports from the CDI questionnaires which showed that these three forms were the most frequently understood amongst the infants across both OWs and CWs. All of the CW stimuli generated mean fixation times of around 50% or lower.

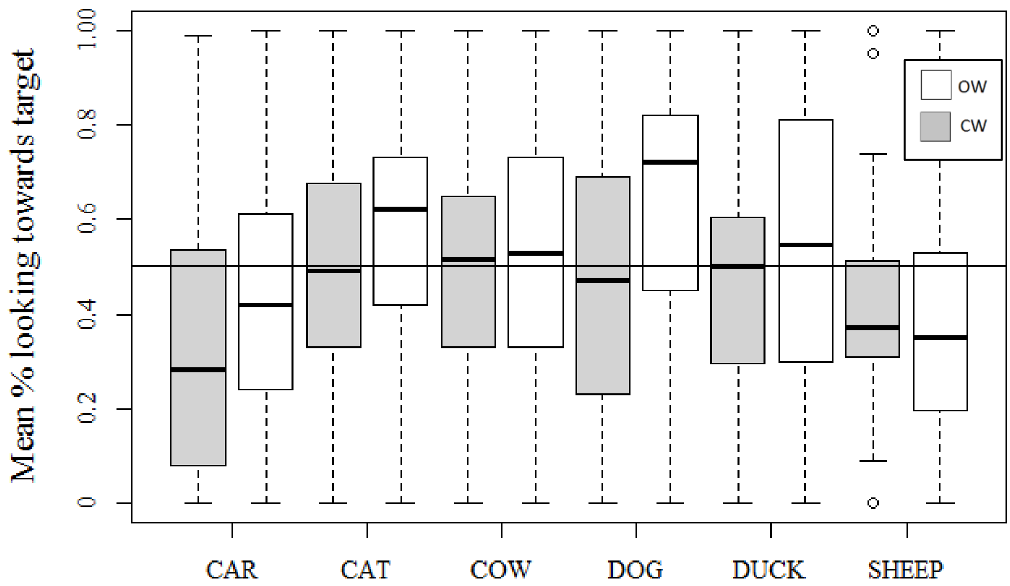

Figure 5.1: Total fixation proportion to target across OW and CW stimuli.

5.3.2 Response latency

An analysis of response latencies across OW and CW forms was then carried out to determine whether there was any difference in infants' speed of recognition across the two conditions. Two infants were excluded from this analysis due to lack of data and anomalously long response times, leaving 20 infants for analysis. Mean response latencies were taken across all OW and CW stimuli for each infant, and Shapiro-Wilk tests confirmed normality for both OW (p=.102) and CW (p=.418) stimuli. A linear mixed-effects model was generated with time to first fixation on the target image as the dependent variable and a fixed effect of condition; item and subject were included as random effects, with by-item and by-subject random slopes for the effect of condition.

No effect was found for condition on infants' response latencies ($\chi 2$ (1) = 1.38, p=.24). On average infants' response latencies were 110ms faster in the OW condition, but as shown in Figure 5.2, variability was high across infants. Furthermore, a Pearson product-moment correlation coefficient indicated a highly significant correlation between infants' responses across the two conditions (n= 20, r = .593, p=.006). This can be seen in Figure 5.3.

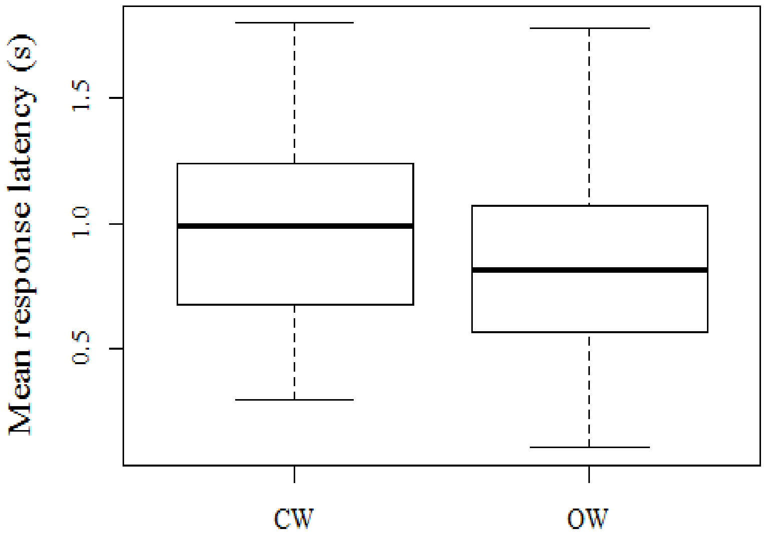

Figure 5.2: Response latencies across stimuli.

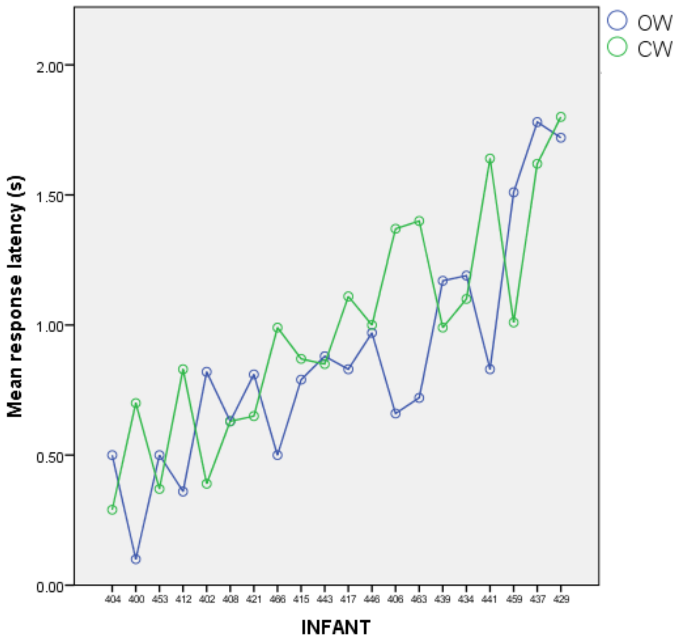

Figure 5.3: Mean response latencies across OW and CW stimuli. Data ordered according to increasing response times.

5.3.3 Language exposure

The CDI questionnaires were then analysed to determine whether infants responded differently to those forms which they heard in the home. Sixteen of the 22 infants were reported to have had regular exposure to OWs, while eighteen of the infants were reported to hear CW forms in the home; 17 of the infants were reported to be exposed to both OWs and CWs, though these were not always corresponding target forms (i.e. some infants were reported to hear *quack* but not *duck*). None of the infants were reported to be able to produce OW forms, while four infants could produce CWs (a combination of one or two out of CAT, DOG and DUCK).

Infants' exposure to the individual stimuli was then controlled for to determine whether their experience of the various forms had any impact on overall responses. This was modelled using a linear mixed-effects regression in R with fixation proportion as the dependent variable. Condition and exposure to each of the twelve stimuli (exposure vs. no exposure) were included as fixed effects, subject and item as random effects, and by-subject and by-item random slopes for the effect of condition. No effect was found for OW exposure (p=.98) or CW exposure (p=.28) on infants' responses. The same model was then carried out with response latency as the dependent variable, and a significant effect was found for exposure: infants looked significantly faster to the target image when they were reported to have been exposed to the target stimuli in the home ($\chi 2$ (1) = 6.38, p=.011). Response latency was shown to increase by an average of 347ms when the infants had experience of the stimuli – this is shown in Figure 5.4. No effect was found for condition here (p=.39).

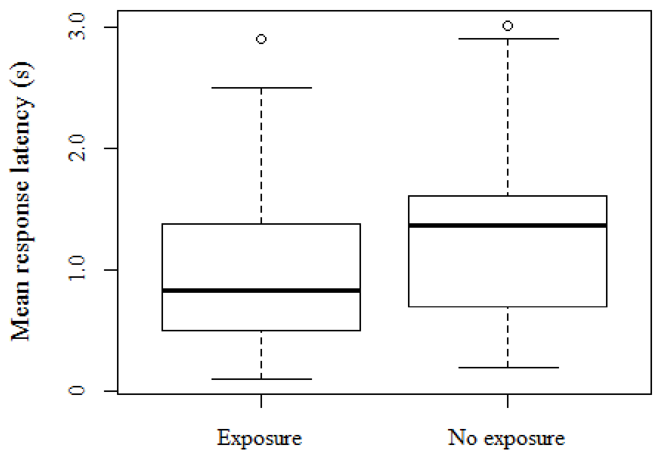

Figure 5.4: Response latency across stimuli according to exposure.

5.4 Discussion

No difference was found between infants' responses to OW and CW stimuli, neither in terms of fixation proportion nor response latency. As also found in Chapter 4, we observed low mean looking across infants: with an average fixation proportion of just above chance in the OW condition and an even lower average in the CW condition, it seems that overall recognition of the stimuli may be unreliable at best. However, infants fixated on the target image significantly faster when they had experience of the stimuli in their input – this was consistent across both OW and CW conditions. The combination of these results allows us to propose that infants do not draw upon any inherent iconic properties in their perception of OWs, and instead we suggest that experience in the input is the main factor affecting their recognition and processing of these forms.

There was a difference of only 50ms between infants' mean response latencies to OWs and CWs across the results, and a distinct trend was observed when we considered the individual infants' response times, which were strikingly similar across OW and CW forms. Again this allows us to refute any position towards an iconic advantage for OWs in early language perception. As in Experiment I, the mean response latencies observed here correspond to Fernald and colleagues' (1998) findings, where infants of 15 months were found to have an average response time of around 995ms. While it is surprising that the younger infants in this study appear to be responding more quickly in both conditions (OW: M=860ms, SD=.43; CW: M=980ms, SD=.43); the wide standard deviations observed across these data can account for this discrepancy. Furthermore, although it could be suggested that the especially fast responses observed in the OW condition may be brought about by the presence of iconicity in these forms, again the lack of difference between responses across stimuli shows that this is not the case: if iconicity were facilitating infants' recognition of OWs, we would expect to see consistent differences across the six OW-CW pairs.

Infants' reported exposure to the individual stimuli had an effect on our results, supporting one of the main hypotheses in this thesis which posits exposure in the input as being a factor in infants' OW acquisition. This is observed in studies by Kauschke and colleagues (2002, 2007), who show that the mothers' use of OWs determines the infants' eventual production of these forms. Evidence was also found in support of parental CDI reports stating that infants understood DOG, DUCK and CAT – as shown in Figure 5.1– but here no clear trend was observed, and the statistical analysis did not support evidence from the CDI reports. However, it is important to note that infants who were reported to hear OWs in the home were in almost all cases exposed to CW forms too. This fits with the infants' responses, and the significant correlation observed across the two conditions. Again this supports a role for the input in these data, as dual exposure to both conditions would explain the lack of effect observed across both fixation proportion and response latency analyses.

We must also consider the infants' ability to map the OW and CW forms to the images presented in the experiment, as only few examples can be found in the literature showing successful sound-meaning mappings amongst infants of such a young age. Bergelson and Swingley (2012) report successful findings from infants as young as six months, but even then the authors recognise that performance is not reliable until infants reach 14 months of age. In a similar eye-tracking design, Mulak and colleagues (2013) found that infants of 15 months were able to map words to their corresponding pictures when the words were produced in a native accent, while Fernald and colleagues (1998, 2006) have used looking-while-listening paradigms to successfully demonstrate word recognition in word-picture mapping tasks at 15 months of age; in both studies, infants' processing improved over the course of the second year. It can therefore be assumed that infants of only 10 months of age will show slower and less reliable responses than those with an extra five months of language experience. Indeed, many similar procedures recruit older participants (Fernald & Hurtado, 2006; Swingley and Aslin, 2000), and Fernald and colleagues (2008) state that infants do not begin to start comprehending words until between the ages of 10 and 14 months. It seems that the optimal age for testing emergent word processing and form-meaning recognition may be 14 months and beyond. Furthermore, in the case of the current procedure, the presentation of words in an unfamiliar voice, paired with the infants' possible lack of exposure to some of the stimuli, may have set them with a task that was too difficult to generate a reliable performance on the group level.

In addition, the majority of the infants who took part in this experiment had not yet begun to speak: sixteen of the 22 infants were reported to produce fewer than three words, and ten infants had not yet produced their very first word form. The onset of word production has been shown to promote young infants' attention to input speech: Majorano and colleagues (2014) observe that infants with increased production abilities show better attentional responses to the sounds that they are able to produce, highlighting how the onset of word production draws infants' attention to the familiar aspects of the input. With these points in mind, it is unsurprising that the results did not show any strong effects owing to the disadvantage that many of the participants faced in terms of their language processing capacities.

5.5 Conclusion

This experiment set out to determine whether the proposed iconicity of onomatopoeia facilitates young infants' ability to recognise OWs over their corresponding CW forms. We observed no advantage for OWs here, and instead exposure to the stimuli in the input was found to be a main factor determining infants' responses. A breakdown of the results showed that infants fixated for longer upon those targets that were most familiar, but again no advantage was found for OWs over CWs; the strongest responses clustered around those targets that were reported to be more familiar to the infants. Furthermore, strong correspondences were found between individual infants' responses to both OW and CW forms, suggesting that recognition of the stimuli was determined by individual infants' experiences and abilities. These results refute any potential advantage for the iconicity of OWs in the early stages of language development, and instead propose a main role for the input. However, we must conclude by acknowledging the problematic nature of these results, as 10-month-olds' ability to successfully take part in a word-picture mapping task such as the one presented here has been called into question.

Together, Experiments I and II provide a new perspective on OW perception and the nature of iconicity in wild and tame forms. However, still the question remains as to whether wildness may play a role in infants' perception and processing of OWs. In the analysis that follows we will address this question more fully, with a consideration of how prosodic features are used by caregivers in the production of OWs when interacting with their young infants.

SECTION III

Onomatopoeia in the input: Infant-caregiver interactions

So far in this thesis we have considered OWs from the infant's perspective, with analyses of production over time and perception in real-time. While the previous four studies have led us to move away from the general assumption that these forms are more easily learned due to their sound-meaning correspondences, our main research question of how or why OWs are acquired in such abundance in the first place remains unanswered. In order to address this question more fully, we must consider the role that OWs play in the early input, as well as the output. Indeed, infants' surrounding linguistic environment is well-established as being of prime importance in shaping their path towards language production.

Evidence of the input's influence can be identified across multiple facets of infant speech from a very young age. De Boysson-Bardies and colleagues (1984), for example, showed that 8-month-olds' ambient language could be identified by adult listeners based solely on the sounds of their babbling. Infants of ten months were also found to produce vowel sounds which differed according to their ambient language (French, English, Cantonese and Algerian Arabic), as they matched the vowel formants produced by adult speakers of those languages (Boysson-Bardies et al., 1989). Similarly, while Levitt and Utman (1992) observed many similarities between the early utterances of one American and one French infant, language-specific features were identified in both the phonemic and the phonological properties of the infants' babble. Prosodic patterns typical of the ambient language have also been identified in French and English infants' prelinguistic vocalisations, as the intonation contours reflected those of the ambient language (Whalen et al., 1991): a predominance of rising contours in French infants' babble and falling contours in English babble. Indeed, the prosodic features of the ambient language are known to play a fundamental role even in infants' very early language perception: in the third trimester of gestation infants' auditory functions are sufficiently developed to enable processing of some aspects of the mother's speech while still in the womb (Vihman, 2014), such that neonates show preferential responses to the prosody of the ambient language over unfamiliar languages at only four days old (Mehler et al., 1988). Over time and with increased experience of the ambient language, infants' linguistic categories are shaped to match those of the surrounding input. Werker and colleagues (1981, 1984) have famously shown that pre-linguistic infants are able to discriminate between contrasting non-native consonants such as Hindi dental /t/ and retroflex /ʈ/ and Salish gottalised velar and uvular plosives /k/ and /q/, even when adults cannot. Furthermore, by the end of their first year, like adults, infants are no longer able to make these distinctions (Werker & Tees, 1984). These results show how language development is shaped by experience, shifting from broad to language-specific perceptual abilities over the course of the first year through increased exposure in the input.

Even when the infant has begun to build the rudiments of a phonological system, still the input determines the words that an infant will acquire: the nature of the input lexicon that infants are exposed to as well as the way in which words are presented are both known to be important determiners of an infant's lexical acquisition. The role of input frequency on infants' language acquisition is discussed in detail by Ambridge and colleagues (2015), who claim that "frequent words are learned before infrequent ones, *all other things being equal*" (p. 243, emphasis in original). The authors discuss the importance of frequency in the input across a number of different linguistic levels (syntactic and morphological as well as lexical), stating that the most frequent forms in the input are acquired earlier and produced more accurately by the infant than less frequent forms. This is supported by findings from Kauschke and Klann-Delius (2007), who show correlations for both word frequency and use of lexical categories between mothers' inputs and their infants' outputs at 1;9 and 3;0. There are of course other important influences on infants' early lexical acquisition, which are acknowledged in Ambridge et al.'s (2015) account. For example, Brent and Siskind (2001) found that mothers' production of words in isolation led to infants' acquisition of those words; use of isolated words in infant-directed speech was more of a predictor of infants' eventual lexical acquisition than overall frequency of occurrence of those words in the input.

The following two chapters will address the question of the input, and how OWs are presented to infants in early language development. A trajectory of development will be observed across the two analyses, allowing us to understand the role that OWs play, first in IDS and then in infant-caregiver interactions, over time. Chapter 6 will begin with an acoustic analysis of speech directed at prelinguistic infants, observing whether there are any prosodic differences in mothers' production of OWs and CWs. This will enable us to determine whether OWs are any more salient than CWs in the input, thus pointing to a potential learning advantage for these forms in early perception. Chapter 7 will then consider how OWs are produced in infant-caregiver interactions over time, not only observing the infants' use of these forms, but also the extent to which they occur in the input, thus enabling us to draw parallels and interactions between the input and the output over time. Together these analyses will provide two different yet complimentary perspectives on infants' experience of OWs in early development: one close-up comparison of OWs and CWs in their abstracted acoustic forms, and one zoomed-out perspective showing the intricate and dynamic reality of OWs in early interactions, as both caregiver and infant speech change in line with the infants' developing linguistic abilities.

6. A prosodic analysis of onomatopoeia in infant-directed speech

6.1 Introduction

Onomatopoeia are commonly reported as a characteristic of 'babytalk', or infant-directed speech (Ferguson, 1964 Fernald & Morikawa, 1993), and could be considered as a lexical feature of this affective speech style which also includes phonological, prosodic and grammatical modifications to speech directed at young infants. Much of the literature focuses on the salient prosodic markers consistently found in infant-directed speech (IDS) as compared with adult-directed speech (Fernald & Simon, 1984; Fernald et al., 1989; McMurray et al., 2013; Stern et al., 1983), but no research to date has studied the combined effects of the prosodic and lexical aspects of this speech register. In this chapter we will explore the extent to which the typical prosodic features of IDS are present in onomatopoeia as produced by mothers in communication with their infants. This analysis will build upon findings from Chapter 4 which showed infants to fixate longer upon hearing stimuli with the most salient features. This will allow us to further consider the nature of iconicity, and whether or not this is a relevant feature of infants' experience of these forms.

Many studies of IDS have found that adults routinely alter the prosodic features of their speech style when addressing young infants; this has been shown to be consistent across both mothers and fathers (Fernald et al., 1989) as well as adults without experience of speaking to infants (Fernald, 1989), and towards infants across a range of ages (Stern et al., 1983). Indeed, IDS appears to be ubiquitous in the early input, and is thought to benefit language development in its early stages through capturing infants' attention (Vihman, 2014) as well as drawing the infant towards specific functional elements of the speech stream (Lee et al., 2008). Lewis (1936) remarks on the "strong affective character" (p.42) of speech directed at young infants, which is first adopted to soothe the neonatal infant, and later to "[make] him smile". More recent empirical research supports Lewis' (1936) claims, as Smith and Trainor (2008) found that infants' positive feedback to IDS reinforces their caregivers' use of higher pitch. Indeed, infants are known to prefer the salient features of IDS over adult-directed speech (ADS, Fernald and Kuhl, 1987), including higher mean pitch, wider pitch range, shorter utterances, longer pauses and repetition (Fernald & Simon, 1984). It has also been shown that caregivers make subtle changes to the typical phonetic categories of the ambient language during the production of IDS (McMurray et al., 2013; Werker et al., 2007), thus simultaneously preserving and emphasising language-specific features in the input[13].

It appears to be unanimously accepted in the literature that IDS is an important and functional aspect of infant language development. Lewis (1936) describes the use of intonation to convey meaning in the absence of linguistic comprehension, stating that the 'affective tone' (p.121) of a word or phrase is what first establishes its meaning, prior to the development of lexical understanding. Even adults have been found to correctly perceive communicative intent through the intonation contours of IDS (Fernald, 1989), demonstrating that "the melody carries the message in speech addressed to infants" (p.1505); in contrast, the intonation of ADS was not sufficient to provide any cues to the information being communicated.

Features of IDS are also claimed to facilitate word segmentation (Golinkoff & Alioto, 1995; Jusczyk et al., 1992), and IDS has been widely proposed as a bootstrapping mechanism for word-learning in young infants. This has been shown for some of the commonly-reported features of IDS: Brent and Siskind (2001) found that infants learned words earlier when their mothers produced them in isolation, while an experiment by Kemler Nelson and colleagues (1989) highlighted 8-month-old infants' preference for pauses at typical clause boundaries when compared with pauses in unnatural positions in a phrase. Golinkoff and Alioto (1995) went some way towards demonstrating the bootstrapping effects of IDS in language learning with their findings on English-speaking adults, who were better able to learn Mandarin Chinese words in IDS than in ADS when they were presented utterance-finally, though target words in utterance-medial position showed no significant effect.

Taken together, this evidence seems to demonstrate a defined role for IDS throughout the language development process, from capturing an infant's attention and

[13] But see McMurray et al. (2013) for a discussion on whether such changes are a feature or a secondary consequence of the modifications typically made in IDS.

communicating affective meaning to providing cues to the lexical and phonological distribution of the ambient language. IDS is thought to facilitate acquisition at all stages of language learning, and it has been found that the characteristics of IDS change as is appropriate to the infant's developing ability (Fernald & Morikawa, 1993). Evidence from the literature demonstrates how specific features of IDS can lead to language learning (Brent & Siskind, 2001; Golinkoff & Alioto, 1995; Keren-Portnoy et al., in press), and so it seems pertinent to relate the use of IDS to features that are commonly found in infants' early lexica. Many studies in this field focus on infants' perceptual preference for IDS (Fernald & Kuhl, 1987; Karzon, 1985; Kemler Nelson et al., 1989), or on typical features of IDS as produced by the caregiver (Lee et al., 2008; McMurray et al., 2013; Werker et al., 2007): while these aspects of IDS are illuminating in themselves, they do not link specifically to the infant's eventual language production. If IDS does direct an infant's word learning, as suggested in the literature, then it should be possible to find further evidence for this correspondence between perception and production of specific word forms in infants' earliest output.

This study analyses the prosodic features of OWs in relation to their corresponding CWs, hypothesising that onomatopoeia are prosodically more salient in IDS than non-onomatopoeic words. Features that are often cited in the literature as being typical of IDS will be examined. It is hypothesised that:

1. Pitch will be modified to result in an increased salience of OWs over CWs: mean pitch will be higher and pitch excursions wider in the production of OWs.
2. Word duration of OWs will be longer than CWs.
3. OWs are produced more frequently than CWs owing to the use of reduplication in these forms (Ferguson, 1983).
4. Pauses will be longer and more frequent before and after the production of OWs than CWs; OWs will appear in isolation more frequently than CWs (Ninio, 1993).

It is assumed that the combination of these features will facilitate the acquisition of OWs in language development, leading to the disproportionate presence of these forms in the early lexicon.

6.2 Method

6.2.1 Participants

Data collected for a previous study was used for this analysis (Keren-Portnoy et al., 2010). Twelve recordings of British mothers interacting with their infants were analysed. Participants were all based in Yorkshire, UK, and were recruited through an advert in a local magazine. At least one parent of each infant held the equivalent of an undergraduate degree from a college or university. The infants, including four females, were all 8 months of age (mean age = 256.6 days), and all had passed a newborn hearing screening; no hearing problems were reported for any of the infants. All infants were either first-born or had no pre-teen siblings.

6.2.2 Apparatus

Data were collected using a Language Environment Analysis (LENA) digital language processor – a recording device placed in a vest worn by the infant. The mother was asked to 'read' with the infant twice daily on each day over a weekend: two picture books selected specifically for the purposes of the original experiment – *Home* (Priddy Books, 2009a) and *Toys* (Priddy Books, 2009b) – were supplied by the experimenters.

6.2.3 Stimuli

The original experiment did not target OWs in any way, and so mothers were not prompted to use onomatopoeia in the book-reading activity: all onomatopoeic words were produced spontaneously. The mothers were asked to talk their infants through each of the books, which presented a series of colourful pictures and their corresponding labels. Importantly, none of the labels used in the books were OWs, though the pictures involved toys and household objects which could elicit onomatopoeic productions from the mothers, including a rubber duck, a train, a car and a jigsaw featuring images of farmyard animals.

6.3 Analysis

OWs and their corresponding CWs produced by mothers during the book-reading task were analysed. A word was considered to be onomatopoeic if it served to imitate the sound of an object in the context of the book-reading task. For example, the mothers used typical OWs such as *meow* to imitate a cat, but also used less typical forms such as *boing* and *bring* to imitate a ball and a bicycle, respectively: in the context of the book-reading task these words were both considered to be onomatopoeic.

Every instance of an OW and its corresponding CW (e.g., *woof* and *dog*, see Table 6.1) were extracted from the recordings using Praat 4.5.02 (Boersma, 2001). Unpaired stimuli, whereby an OW was produced without the production of at least one corresponding CW in the same recording, and vice versa (for example, *quack* occurring without *duck*, or *ball* occurring without *boing*), were excluded from the analysis, in order to ensure that pairwise comparisons could be made across OW and CW forms for each mother. Wherever both OW and CW forms appeared in the same recording, whether they occurred together or in separate contexts, they were considered a pair. The set of OW-CW pairings included in the study is detailed in Table 6.1, along with the stimulus name for each pairing, shown in SMALL CAPITALS.

Table 6.1: Stimuli and variables used in the analysis.

Stimulus	OW	CW
BALL	Bounce/Bouncy/Boing	Ball
BEE	Buzz	Bee
BICYCLE	Bring bring (of bell)	Bicycle
CAR	Brum/Vroom	Car
CAT	Meow	Cat
COW	Moo	Cow
DOG	Woof	Dog
DUCK	Quack	Duck(ie)
FROG	Ribbit	Frog
HORSE	Neigh	Horse
PIG	Oink	Pig
SHEEP	Baa	Sheep
TRAIN	Choo choo/Toot toot /Woo woo	Train
TELEPHONE	Ring ring	Telephone

As is typical in IDS, many instances of OWs were reduplicated in the recordings (e.g. *woof woof*, Sundberg, 1998). With this in mind, reduplicated OWs were analysed as single units in cases where there was a pause of less than 200ms between tokens, while pauses of more than 200ms marked a new token even in cases of multiple reduplication. This is shown in example (1), where numbers in brackets indicate pause duration (in seconds):

(1)
```
M1| it's a duck (3.45)
M2| quack quack (2.32) quack quack (2.12)
```

Although the token *quack* occurs four times in this example, for the purposes of this analysis this counts as two tokens of *quack*, each with an instance of reduplication. This approach takes into account the repetitive characteristics of established onomatopoeic sequences (e.g. *quack quack*, *woof woof*), while also acknowledging reduplication as a typical feature of infant-directed speech (Sundberg, 1998). On a methodological level this also makes for a more accurate picture of word duration, as the inclusion of long pauses between tokens could create a bias towards OWs.

Praat was used to measure mean pitch, pitch range and duration for each of the stimuli, as well as pauses separating the stimuli from surrounding speech. Measurements were taken from word onset to offset, including aspiration of word-final consonants where appropriate. Pitch traces were cross-checked to ensure that they corresponded to the audio data, and any errors were corrected manually in Praat. Measurements for every individual OW and CW token were recorded. Transcriptions were also made of the utterances containing the OWs and CWs used in this analysis in order to account for word isolation. As in Brent and Siskind's (2001) analysis, words were considered to be fully isolated if they were separated from other words in the speech stream by a pause of at least 300ms on both sides of the target word. Partially-isolated words were considered if they had a pause of 300ms either preceding or following the target.

6.4 Results

6.4.1 OW production across mothers

On average, 20 minutes and 12 seconds of recording were available for each mother (min= 5 minutes 25 seconds, max = 40 minutes, 20 seconds) from the book-reading task, from a total of 31 separate recordings (mean= 2.58 recordings per mother). The mother with the shortest recording produced 8 OWs in total and 10 corresponding CWs, while the mother with the longest recording produced 17 OWs and 39 CWs. Given the difference in recording time of almost 35 minutes across mothers, a Pearson product-moment correlation coefficient was used to analyse the distribution of OWs in the data; this indicated that there was no correlation between duration of recording and number of OWs produced by the mothers (r= .012, n= 12, p=.971).

The frequency of production of each OW and CW is detailed in Table 6.2. As shown here, production was almost equal across OWs and CWs, in terms of the number of mothers that produced each of the forms and the number of times they produced them. While the use of OWs was highly variable across different mothers, all of the mothers produced at least two of the OW-CW pairs listed in Table 6.1 (max = 11, min= 2, mean= 5.17). Furthermore, seven of the twelve mothers produced at least five of the pairs, providing a wide pool of stimuli for comparison. A Shapiro-Wilk test confirmed normality for word duration and mean pitch for both OW and CW stimuli across mothers (word duration: OW p=.289, CW p=.506; mean pitch: OW p=.169, CW p=.735), as well as pitch range of CWs (p=.735), but this was not normally distributed for OWs (p=.014). These data (both OW and CW tokens) were thus normalised using a log10 transformation. Parametric tests will be used for all analyses; by-subject random slopes will be included in all statistical models, but by-item random slopes will be omitted here, since each mother produces a different set of OW-CW pairs[14]. All figures represent the original non-transformed data.

[14] With thanks to C. Scheepers (personal communication) for advice on this issue.

Table 6.2: Frequency of OW and CW production in the dataset.

Stimulus	OW		CW	
	mothers	tokens	mothers	tokens
BALL	6	13	7	28
BEE	6	11	6	12
BICYCLE	1	2	1	2
CAR	7	27	7	23
CAT	8	11	8	10
COW	1	2	1	1
DOG	5	8	5	8
DUCK	11	95	11	95
FROG	1	3	1	2
HORSE	4	5	4	6
PIG	1	1	1	1
SHEEP	1	1	1	1
TRAIN	9	35	11	35
TELEPHONE	2	2	2	2
TOTAL		216		226

6.4.2 Pitch

A linear mixed-effects model compared the mothers' pitch across OW and CW stimuli, with mean f0 as the dependent variable. Type of stimuli (OW or CW) was included as a fixed effect, with subject and item (target word) as random effects and by-subject random slopes for the effect of type of stimuli. Stimuli type had a significant impact on the production of the target word ($\chi 2$ (1) = 4.507, p=.034), as OWs had an increased mean pitch of about 65Hz (see Figure 6.1).

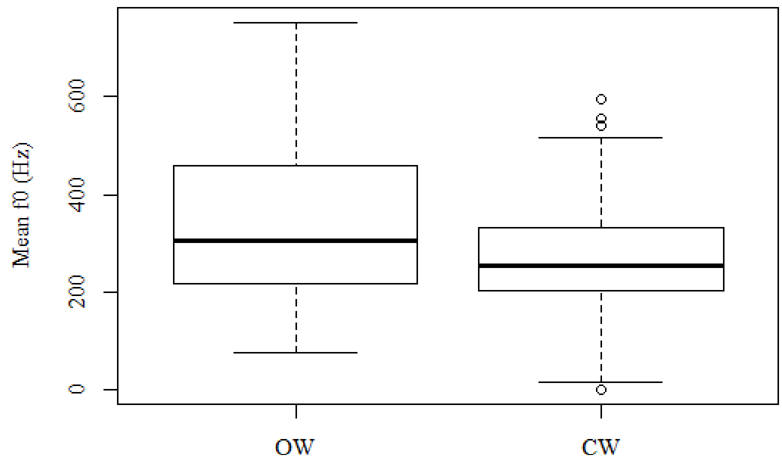

Figure 6.1: Mean pitch values for OWs and CWs produced across mothers.

Pitch range was then compared across OW and CW stimuli, and OWs were found to be produced with a significantly wider pitch range ($\chi 2$ (1) = 5.32, p=.021), with an average increase of around 30.5Hz in the OW condition (see Figure 6.2). These two results are in line with hypothesis 1: in terms of the typically-reported pitch features of IDS, OWs were found to be more salient than their CW equivalents.

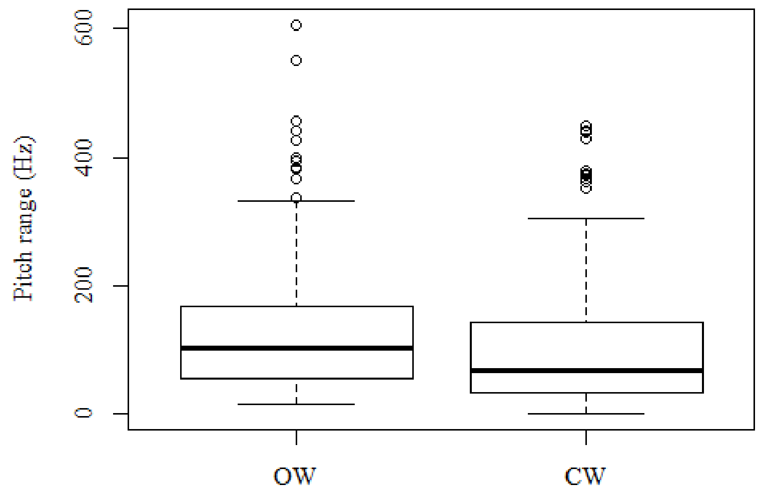

Figure 6.2: Mean pitch range for OWs and CWs produced across mothers.

6.4.3 Word duration

It was expected that OWs would be longer than their respective CWs, due to the fact that OWs are commonly produced with reduplication (e.g. *quack quack*). Indeed, of the 216 instances of OWs in the dataset, 84% (n=181) were reduplicated, with all but two instances undergoing full reduplication. Reduplication did not occur in any of the CWs in the dataset. While there were some cases of extensive reduplication across tokens (for example, OW BEE was reduplicated 25 times in one instance), the vast majority of OWs (71%) were reduplicated twice. CAT and HORSE were the only two OWs to feature no reduplication across the full dataset, while DOG and BALL were the only OWs which were always reduplicated.

A linear mixed-effects model compared word duration across OWs and CWs, with word duration as the dependent variable, type of stimuli as a fixed effect, and subject and item as random effects. By-subject random slopes were included for the effect of type of stimuli. OWs were found to be significantly longer in duration than CWs ($\chi 2$ (1) = 15.165, p<.000); mean duration values show the OW stimuli to be 659ms longer than CW stimuli on average, but as shown in Figure 6.3, there is wide variability in OW duration. This confirms hypothesis 2, again showing that OWs are more salient than their equivalent CWs. However, it is not clear whether this extended word duration is prompted by reduplication or by vowel or consonant lengthening.

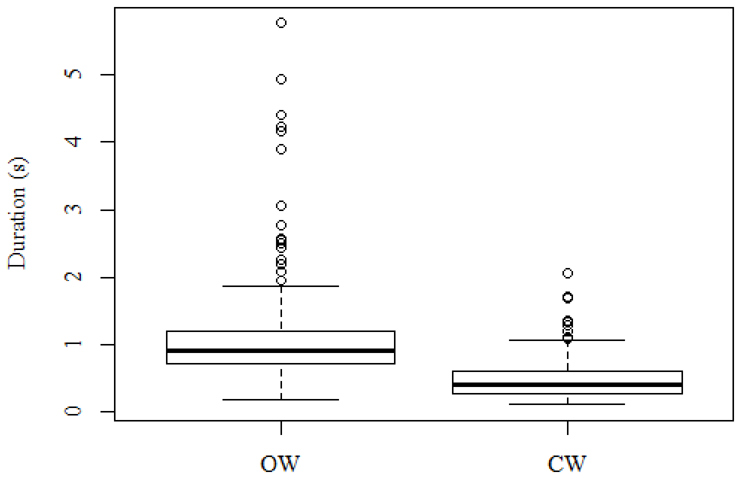

Figure 6.3: Mean word duration across OWs and CWs.

An exploratory analysis then considered OWs separately to observe whether the presence of reduplication had any effect on the duration of these forms. A linear mixed-effects model with word duration as the dependent variable and reduplication as a fixed effect (including subject and item as random effects and by-subject random slopes) showed a near-significant effect for reduplication on the duration of OWs ($\chi 2$ (1) = 3.657, p=.056); reduplicated OWs were on average around 402ms longer than non-reduplicated forms. This suggests that both reduplication and vowel/consonant lengthening may be relevant to the longer duration of OWs.

Finally, it was proposed that the observed increase in pitch range of OWs may be prompted by their longer duration. A Pearson product-moment correlation coefficient revealed a highly significant correlation between pitch range and word duration across all OW and CW tokens in the dataset (r= .251, n= 444, p<.000). In order to account for this, rate of pitch change (y) was calculated across all targets with the equation $y = \frac{pitch\ range\ (Hz)}{duration\ (ms)}$; this takes into consideration the change in pitch across a word in terms of its duration. A Shapiro-Wilks calculation showed a non-normal distribution for rate of pitch change across OWs (p<.000), and so this measure was normalised in R using a log10 transformation. A linear mixed-effects model with rate of pitch change as the dependent variable and type of stimuli as a fixed effect was then carried out. Subject and item were included as random effects with by-subject random slopes for the effect of type of stimuli. Results showed a significant difference between OW and CW production ($\chi 2$ (1) = 7.375, p=.007); rate of pitch change was significantly higher across CWs than OWs by around 400Hz/second. This contradicts the trends observed above in the analysis of absolute pitch range, as it suggests that the wider pitch excursions noted in OWs may only occur as a function of their extended duration.

6.4.4 Repetition

It was proposed in hypothesis 3 that OWs may occur more often than CWs in IDS owing to the presence of reduplication. However, as observed in example (1), the presence of repetition alongside reduplication may also be a factor in the mothers' OW production. Repetition was thus considered alongside reduplication in order to account more thoroughly for any frequency effects. The definition of reduplication used here (see above) does not account for the extent to which OWs are repeated in full within close temporal proximity. Fifty eight percent (n= 126) of the OWs produced in the dataset – both reduplicated 'clusters' such as *woof woof* as well as those without reduplication such as *meow* – are repeated in immediate proximity to another token of the same OW (with or without reduplication), separated only by a pause. Furthermore, 87% of all OWs in the dataset occur with either reduplication or immediate repetition: essentially nearly all OWs occur directly next to another instance of the same word. Importantly, 45% of OWs are both reduplicated *and* repeated within the same utterance (see example (1), M2, above), thus providing multiple tokens of the same word type, one after the other. In contrast to this, only one instance of direct repetition can be found across all 226 CWs, and there are no reduplicated CWs in the dataset.

A generalised linear mixed-effects model was generated using the *glmer()* function in R to account for the binomial distribution of this data (repeated vs. non-repeated). Use of repetition was included as the dependent variable, with type of stimuli as the fixed effect, subject and item as random effects and by-subject random slopes. Unsurprisingly, repetition featured significantly more often in OW production ($\chi 2$ (1) = 28.61, p<.000). However, post-hoc paired-samples t-tests comparing the proportion of repeated vs. non-repeated OWs across mothers showed no difference between the proportion of non-repeated (M=50.07, SD=25.1) versus repeated OWs (M=49.9, SD=25.1; t(11) = -.01, p=.992). When reduplication and repetition were considered together, the proportion of repetition or reduplication in the production of mothers' OWs (M=87.33, SD=13.6) was significantly higher than the proportion of OWs with neither of these features (M=12.67, SD=13.6; t(11) = 9.511, p<.000).

Repetition was then considered in terms of the mean pitch, pitch range, rate of pitch range and duration of OWs, to determine whether the extensive use of OW repetition brought about any prosodic changes in the mothers' production of these forms. Four linear mixed-effects models considering the OW data only were carried out in R, with mean pitch, pitch range, rate of pitch change and word duration as the four dependent variables, each with repetition as the fixed effect (repeated vs. non-repeated) and target word and subject as random effects. By-subject random slopes were also included. No effect was found for any of the four measures (mean pitch: p=.36, pitch range: p=.41 , rate of pitch range: p=.48, word duration: p=.84).

6.4.5 Isolated words

Pauses before and after all OWs and CWs in the dataset were analysed to account for fully isolated (pauses before *and* after the word is produced) and partially isolated words (pauses *either* before *or* after the word is produced). As detailed above, a pause was considered for analysis if it measured 300ms or more in duration, in line with Brent and Siskind's (2001) approach.

In agreement with hypothesis 4, the analysis shows that OWs occur in isolation more often than CWs: 53% (n= 114) of OWs produced in the dataset appeared in full isolation, while only 5% of CWs (n= 11) were fully isolated. Unsurprisingly, a generalised linear mixed-effects model with isolation (isolated vs. non-isolated) as a dependent variable and type of stimuli as the fixed effect showed that OWs were produced in isolation significantly more often than CWs ($\chi2$ (1) = 15.306, p<.000). When partial isolation in the mothers' production of OWs was also considered, a further 94 OWs (44%) were found to be partially isolated, with a pause either preceding or following the word. The same generalised linear mixed-effects model, this time with the inclusion of partial as well as full isolation in the dependent variable (full or partial isolation vs. no isolation) again showed OWs to be produced significantly more often in full or partial isolation than CWs ($\chi2$ (1) = 26.722, p<.000). In total 97% of OWs were produced in at least partial isolation compared with 44% of CWs. Figure 6.4 shows the percentage distribution of use in isolation across OWs and CWs.

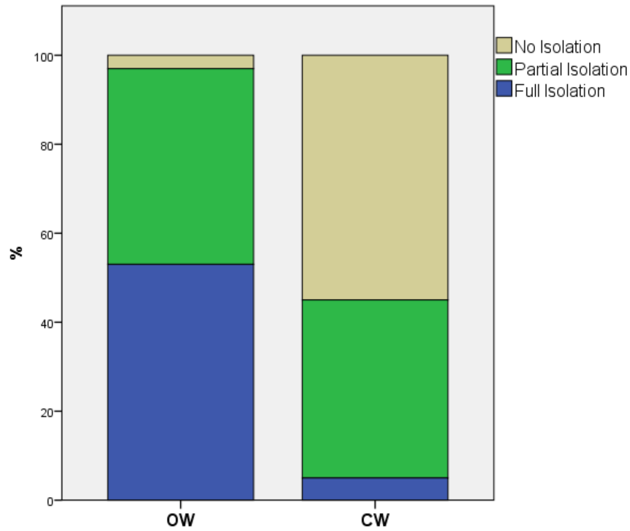

Figure 6.4: Percentage distribution of use of isolation across OWs and CWs

The distribution of word-initial and word-final pauses can be accounted for when we consider the trends in OW and CW production that are observed throughout the data. A breakdown of these pause types showed word-final pauses to be more common following CWs than OWs: on average, 44% of all CWs are produced with a word-final pause, compared with 23.5% of OWs. This trend can be attributed to a specific speech-style that the mothers use in addressing their infants, whereby both OWs and CWs are produced within syntactic 'frames'. Some typical examples can be seen in (2) to (4) (CWs are highlighted in bold):

(2) *Joshua*

```
M1| a buzzy bee (.26) bzbzbzbzbzbz (.79)
M2| and a duck (.69) quack quack (.69) quack quack (1.69)
M3| and a cat (.49) meow (1.31)
M4| and a dog
```

(3) *Lily*

```
M1| that's a duck (.51) quack quack (.27)
M2| and a sheep (.19) baa (.52)
M3| s'a pig (.22) oink oink (.82)
M4| s'a cow (.63) moo (.81) moo (1.59)
M5| there's a bowl
```

(4) *Warren*

```
M1| is that a duck (.41) quack quack quack (.76)
M2| quack quack (.76) quack quack (3.6)
M3| it's a bicycle (1.83)
M4| bicycle (.16) bring bring (.) bring bring (.)
M5| bring bring (.57) there's a
```

As shown here, all three mothers use the same syntactic structure when engaging with their infant in the picture book-reading activity. Word-final pauses appear to be common across CWs, as they occur after a repeated existential phrase ('there's a', 'and a', '[it]'s a') and are followed by a corresponding OW, which is produced in isolation on the back of the word-final pause. Furthermore, all three examples show the use of reduplication and repetition of the OW, whereas (4) is the only example containing repetition of a CW, which in this instance is produced in isolation – the only instance of direct CW repetition in the dataset. While our primary aim is to consider the prosodic features of OW production, the apparent syntactic patterning of OWs and CWs as shown in these examples may be an important feature of OW production in IDS. As such, the distribution of OWs and CWs on a syntactic level will now be considered.

6.4.6 Proximity

Following the analysis of OWs and CWs produced in isolation we considered the proximity of an OW to its corresponding CW. As shown in examples (2) to (4) above, there appeares to be a pattern here: in many cases the mothers produced CWs in immediate proximity to their corresponding OWs. If consistent across the dataset, an analysis of OW-CW proximity might provide an important insight into the use of OWs in IDS.

A 'proximity score' was calculated from every OW to its nearest corresponding CW, whereby the number of words produced between the OW and the CW was counted for each OW in the dataset (for example 'a **train** that goes **choo choo**' would have a proximity score of 2, as there are two words between the OW and the CW). As some CWs were produced in a context without the OW counterpart in close proximity (but not vice versa), the initial analysis was based on OW rather than CW production.

Of the 216 OWs analysed in the full dataset, 194 (90%) were found to occur within 10 words of the corresponding CW (M= 0.77 words), and over half (n= 127) were produced immediately next to the corresponding CW. Again this shows evidence for a routinized approach to OW production: these forms appear to depend on the presence of a CW. When the analysis is reversed to consider the proximity of OWs to CWs, the figures are less illuminating but still show the same trends. Seventy four percent of CWs are produced within a 10-word proximity to a corresponding OW (M= 1.6 words), and 81 of these (36% of all CWs in the dataset) occurred immediately next to the OW in the mothers' speech. Here we see that CWs do not necessarily occur with their corresponding OW, but nevertheless mothers produce the accompanying OW form in the majority of cases.

6.5 Discussion

This analysis has shown that mothers' production of OWs in IDS is more salient across the board than their production of the corresponding CWs. In terms of the common features of IDS (Brent & Siskind, 2001; Fernald & Kuhl, 1987), OWs are more prominent in the input than their CW counterparts with regard to pitch, pitch range and word duration. Reduplication and repetition are significantly more frequent in the production of OWs than CWs, as are occurrences in isolation. Proximity of OW-CW pairings was also found to be an important feature of OW production, as OWs occurred almost exclusively in close proximity to – often immediately next to – their CW counterpart.

The analysis of pitch range was confounded slightly when it was combined with the duration of the word in question, as the extended word length of OWs appeared to provide more opportunity for wider pitch excursions. Indeed, when duration was controlled for, rate of pitch change was higher in the CW forms. This demonstrates the dynamic effect of production on prosody, which was found here to be dependent on multiple factors and not only on the lexical status of the word in question. However, considering the infant's experience of OWs, absolute pitch range may be a more appropriate measure to adopt in the analysis of IDS, since the combination of longer words and wider pitch excursions in the production of OWs undoubtedly serves to increase the salience of these forms.

Since the production of OWs involves the stylised imitation of non-human sounds, we must consider the nature of the prosodic effects in terms of individual word forms. Indeed, a wider pitch range may not be appropriate for some OWs. As shown in Figures 6.5a-c, a particular pitch may be implicit in the production of a specific OW, such as monotonal high-pitched *bring bring* (TELEPHONE) compared with a rising variable pitch in *ribbit* (FROG) or a falling variable pitch in *neigh* (HORSE): here we see pitch being used variably to represent the OW in question.

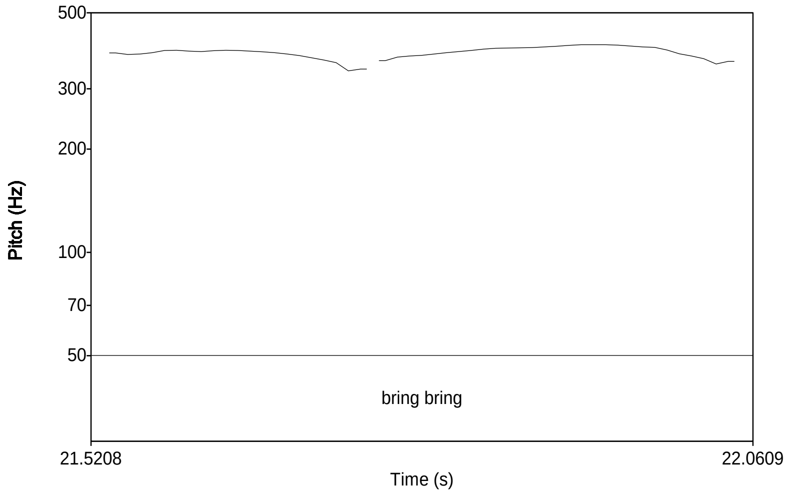

Figure 6.5a: Pitch trace of OW BICYCLE produced in IDS.

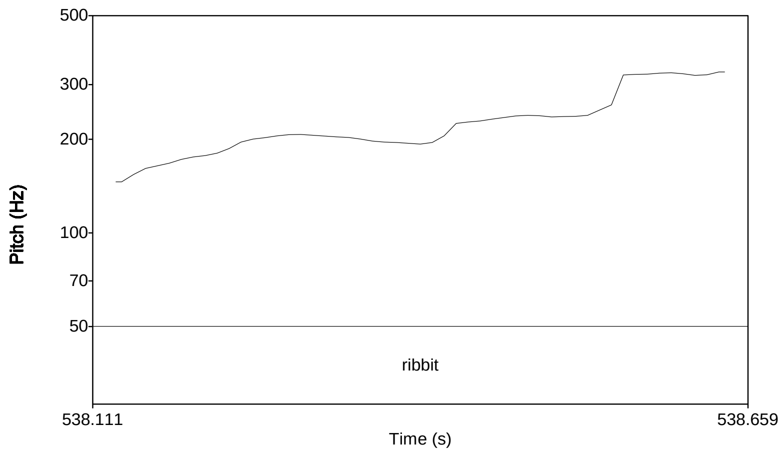

Figure 6.5b: Pitch trace of OW FROG produced in IDS.

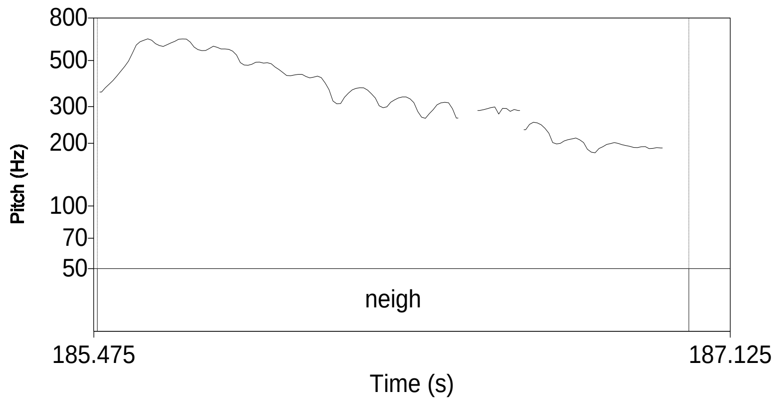

Figure 6.5c: Pitch trace of OW HORSE produced in IDS.

Longer word duration was largely attributed to the presence of reduplication in OWs. However, reduplication was not consistent across all stimuli: no instances of CAT or HORSE were reduplicated, and these targets still exhibited longer word length in the OW than in the CW. It seems that two important features of OWs are at play here, one which is among the commonly reported characteristics of IDS – increased word duration, which applies to an even greater extent in OW than in CW production; and one which is a typical feature of onomatopoeia in general – reduplication, which enhances the salience of these forms through the reiteration of specific segments in the input. We can relate this evidence to the developmental literature, as the effect of frequency of particular forms and structures in the input is known to be an important factor in an infant's language learning (Ambridge et al., 2015): the use of reduplication in the production of OWs effectively doubles the number of OWs presented to the infant, thus giving these forms a distinct learning advantage over their non-reduplicated CW equivalents. Together, the use of repetition *and* reduplication in the production of OWs brings about their increased presence in the input; repetition is cited as one of the typically salient features of IDS (Brent & Siskind, 2001; Fernald & Kuhl, 1987), yet there was only one example of CW repetition in the entire dataset.

Taken together, these results provide a new perspective on onomatopoeia in early language production, suggesting that the salience of these forms in the early input may drive their early production in the output. This calls into question Werner and Kaplan's (1963) theory of implicitly learnable sound-meaning correspondences in onomatopoeia, as well as claims made by Imai and Kita's (2014) 'sound symbolism bootstrapping hypothesis': OWs do indeed appear to provide a perceptual advantage to infants in the early stages of word learning, but this may not be implicitly specified by any intrinsic sound-meaning correspondences – rather, it is explicitly determined by the prosodic nature of the caregiver's OW production. Indeed, Werner and Kaplan's review skims over the role of the input, thus neglecting to consider infants' early experience of language: Leopold's (1939) account of his daughter's language development is cited widely in Werner and Kaplan's analysis, yet they do not acknowledge the author's descriptions of his daughter's input, notably with regard to onomatopoeia. While the proposal that infants are more easily able to connect sound and meaning in words such as onomatopoeia may be theoretically appealing, this approach disregards the reality of language learning as an interactive process, incorporating both perception and production from the outset. We can also consider previous findings from this thesis in light of these results, as in Chapter 4 infants showed a processing advantage for the stimuli with the highest pitch. As these features have been found here to be typical of OW production in IDS, it seems that infants' responses in Chapter 4 may be typical of OW perception in general. Indeed, the processing advantage that was identified for the most salient forms in our eye-tracking study may extend to OWs in caregiver speech.

With this in mind, it is perhaps to be expected that infants may acquire so many OWs in the early lexicon, as we can establish a functional role for many of the features analysed in this study in terms of evidence from the wider IDS literature. As Fernald and Kuhl (1987) show, young infants tend to prefer the exaggerated pitch contours of IDS, which have been found to attract attention more reliably than the pitch features found in adult-directed speech (Fernald, 1985). On this basis it can be assumed that the further increase in salience of OWs in terms of mean pitch and perhaps also pitch range draws infants' attention to these forms over the less-salient CWs. Further to this, Brent and Siskind (2001) demonstrate that mothers' production of isolated words in IDS impacts directly upon their infants' eventual word production, due to the facilitation of segmentation through the framing of words with pauses. This is supported by recent findings from Keren-Portnoy and colleagues (in press), who show that prelinguistic infants are better able to recognise novel words that they have heard in isolation in the early input over words which have been embedded in sentences.

Finally, a revealing trend was found in the syntactic distribution of OWs: while CWs were commonly found without an OW equivalent in close proximity, the production of OWs without the corresponding CW was rare. Again, this demonstrates a specific approach in mothers' production of OWs, suggesting a non-random or conventionalised use for OWs in IDS, which may help infants to bootstrap into language learning. Indeed, a striking similarity was observed across the mothers' production of OWs, which often appeared with their corresponding CWs in syntactic 'frames'. Evidence from the literature suggests that this is also beneficial to infants' language learning, and may support the acquisition of CWs, as well as OWs. Mintz (2003) posits that the use of 'frequent frames' in IDS facilitates infants' processing of word categories, making it easier for infants to compute running speech, and to categorise lexical and grammatical information in the input.

Werner and Kaplan's (1963) position on increasing differentiation can also be reconsidered in light of these findings. The proximity analysis showed how OW and CW forms occur in close proximity to one another in almost all instances. The production of OW-CW combinations in infants' early words can thus be accounted for by the consistent pairing of these word forms in the input. Here we find empirical evidence for the OW-CW combinations that are commonly found in infant speech being modelled in the input, as opposed to Werner and Kaplan's exclusively theoretical proposition positing this phenomenon as an idiosyncratic child invention which develops ontogenetically over time as the infant moves towards the more abstract lexical form.

We must consider why OWs might lend themselves to a more salient prosody than their CW equivalents. One reason may be their distance from the adult language, as their marginal role in adult speech confines them to the 'babytalk' lexicon, thus enabling more flexibility in their production. Indeed, the fact that these forms are perhaps superfluous to the adult language could be advantageous in IDS, since the prosodic conventions that normally govern adult-directed speech do not apply to OWs. This can be seen in the wide prosodic contrasts between OW and CW production here. The consistent OW-CW pairings also point to the status of OWs in the adult language, as caregivers may not consider these forms to be words in their own right. This might also explain their predominant use in isolation, as onomatopoeia do not have a specified grammatical role, serving instead as embellishments to an appropriate phrase or word form. In interactions with 8-month-olds, where the infant typically cannot respond verbally to the input, OWs provide caregivers with lexical variety which may serve to engage the infant more successfully than CWs. Positive infant engagement has been found to reinforce mothers' use of higher pitch contours in IDS (Smith & Trainor, 2008), and it seems that the use of OWs in this study may have provided caregivers with material for joint engagement with their infants.

It could also be proposed that the pairing of OWs with CWs in the mothers' speech may reflect attempts to maintain the infants' attention during the experimental task, as OWs may serve as a more reliable way of involving infants in such an activity. Indeed, infants' responses to the task during the data collection demonstrate anecdotally their engagement with the mother as she produced OWs: many of the infants made noises and cries of excitement during the mothers' production of OWs, and one infant even produced the word *quack* when the mother was talking about the picture of the duck – the only comprehensible word produced by any of the infants in these recordings. This brings us back to the findings of Kauschke and colleagues (2002, 2007), who acknowledge the "attention-getting" function of OWs, which serves to promote "involvement in conversation" (2007, p.198). When this acoustic analysis is paired with Kauschke and colleagues' (2007) lexical analysis of IDS it is possible to draw solid conclusions regarding the nature of OWs in the infants' input and its influence on their eventual output.

6.6 Conclusion

This study has demonstrated a revealing connection between onomatopoeia and IDS. These findings, together with other recent studies, lead to the conclusion that the common production of onomatopoeia in infants' early words is brought about by the input received in IDS, rather than by any features intrinsic to the words themselves. Findings revealed a tendency amongst caregivers to produce onomatopoeia in a similar prosodic and syntactic style across the board: it seems that onomatopoeia may not only be conventionalised in terms of the infant input – that is, producing the form *choo choo* to describe a train, which almost certainly has no correspondence with the noise that a train may make – but also in terms of the adult output, as the mothers were all found to apply the same routines and features to onomatopoeia throughout the recordings. Claims from the literature which propose a symbolic advantage for these forms in infant language learning appear to be far removed from the reality of infants' production of conventionalised onomatopoeia: these forms are no less arbitrary than their conventional equivalents, but are made more salient and appealing (and thus more learnable) in the input through the use of prosodic features that are particular to IDS. It seems that onomatopoeia may occur so prominently in early language development owing to their position – both prosodic and lexical – in IDS. Indeed, their presence in early infant speech appears to be a product of the affective linguistic mechanisms that are unconsciously but effectively put into practice in the adult output.

7. OW production in caregiver-infant interactions

7.1 Introduction

In Chapter 6 we observed how mothers made use of routinized syntactic structures when producing OWs in interactions with their infants. It was suggested that the use of 'frequent frames' (Mintz, 2003) in the presentation of OWs (and CWs) may be of benefit to the infants' segmentation of these forms, in terms of the development of both lexical and grammatical categories. However, as Kauschke and Klann-Delius (2007) observe, there may also be a role for OWs with regard to "involvement in conversations" (p.198), suggesting that the use of these forms in IDS may have an interactive function in the early input. It was posited that the presentation of CWs and OWs within routinized syntactic structures demonstrated an important source of early joint engagement, as OWs appeared to be used functionally to maintain the infants' attention during a book-reading task. Indeed, if we return to example (2) from Chapter 6 (reproduced in (1) below), we see how the prosody and structure of these 'frames' may be used to imitate conversational roles, even before the infant is able to speak:

(1) *Lily*

```
M1 | that's a duck (.51) quack quack (.27)
M2 | and a sheep (.19) baa (.52)
M3 | s'a pig (.22) oink oink (.82)
M4 | s'a cow (.63) moo (.81) moo (1.59)
```

Although the infants were unable to contribute verbally in response to these interactions, still a distinct 'turn-taking' dialogue can be observed, with the mothers appearing to fill the role of both interlocutors. As shown in examples (2) to (4) (Chapter 6), pauses occur between the production of the CW and the OW, which could be interpreted as the mother's implementation of a 'question and response' routine. This demonstrates how OWs can be used to engage infants in interactions which imitate turn-taking dialogues, with the use of rhythmic and prosodic features to simulate the characteristics of adult conversation.

Jaffe and colleagues (2001) show mother-infant dialogues to be bidirectional, even with infants as young as four months. Indeed, features such as the timing of turn-taking, overlap and pause duration between turns is present in dialogues with infants who have not yet begun to speak. This supports the suggestion that the routines observed in the mothers' production in Chapter 6 may have constituted more than simply one-sided monologues, and that the infants may have been more than passive observers to the routines that were taking place. While no evidence towards this was identified in the recordings (i.e. the infants did not appear to 'take part' in these interactions through vocalisation), visual evidence might have revealed a more active engagement on the infants' part, perhaps through pointing, gaze-shifts or the use of gesture.

In a review of dialogues in adult speech, Garrod and Pickering (2004) discuss the importance of repetitions in turn-taking routines, which simplify interactions for both interlocutors through the reuse of linguistic information, as well as facilitating language processing via the avoidance of ambiguity. This brings about the "interactive alignment of linguistic representations" (p.11), which may be especially important in early language development. The authors maintain that language learning takes place through dialogue, suggesting that early interactions play an important role in an infant's acquisition of language. Indeed, caregiver-infant interactions are known to be essential to language development, and the use of joint attention is well established in the literature as being an important determiner of an infant's later language ability (Saxon, 1997; Tomasello, 2003). Bee and colleagues (1982) found that the quality of mother-infant interactions in the first year of life influenced the infants' IQ as well as language proficiency at four years of age, and the same can be said for infants with both minimal and more established language experience: parents' feedback to prelinguistic infants' babbling has been found to prompt further phonological learning (Goldstein & Schwade, 2008), while Hargrave and Sénéchal (2000) showed that even at five years of age, infants with language delays experience improvements in their vocabulary through joint engagement in book-reading tasks.

Zimmerman and colleagues (2009) propose that language development is brought about by early conversations: infants aged 2-48 months were observed over an 18-month period, and the authors found that those infants who were engaged in more conversations with the adult showed better language skills on standardised tests. Furthermore, adult-child interactions are posited as supporting the caregiver's sensitivity to their infant's developing language abilities, prompting them to engage in a way that promotes word production while also scaffolding further language learning. This is in line with Vygotsky's (1978) 'zone of proximal development' (ZPD), which relates to the difference between a child or infant's independent abilities in relation to their abilities when guided by an adult. That is, support from an adult may increase an infant's ability to perform a task through the provision of "natural scaffolding of the ZPD" (H. Cain, personal communication). As Vygotsky explains, "the zone of proximal development defines those functions that have not yet matured but are in the process of maturation" (1978, p.86): this relates to the infant's own internal "dynamic developmental state" (p.87), but is dependent on interactions with caregivers in the surrounding environment. Vygotsky posits that an infant's development is always lagging behind any learning processes: development advances when a new word is learnt, for example, but subsequently the infant's learning capacity moves forwards at this point, which in turn will drive forwards their overall language development.

With this in mind, it could be posited that the simulation of turn-taking 'dialogues' with the routinized presentation of OWs, as seen in Chapter 6, may support infants' early acquisition of these forms as well as their understanding of conversational routines. However, the limitations of the data collection make it impossible to understand the nature of these interactions, as we do not have access to the longitudinal data that might demonstrate the infants' developing role in these turn-taking routines. This chapter will build upon observations from Chapter 6 by using video data to explore these early interactions in more detail, examining how they change as the infant goes from rudimentary word production to more developed linguistic abilities. This will provide a perspective on the social and interactional role that OWs play in early language development, and how the infant's feedback shapes the use of OWs in IDS.

As well as observing the nature of OWs in turn-taking routines, the present chapter also aims to bring together the findings from the previous analyses in this thesis, with a broader view on the dynamic interaction between perception and production which may bring about the use of OWs in early development. OW production by both infants and caregivers will be observed over time, relating to the findings from Chapters 2 and 3 showing the shift from OWs to CWs in infants' early words. Observations of how OW and CW production takes place in real-time will also bring to light the role that these forms play in early development, in relation to both the individual infant's language learning and their understanding of communication as a social tool. An exploration of the nature of OWs in turn-taking interactions will provide an insight into their role in infant-caregiver communication, supporting evidence from Chapter 6 by showing the extent to which these forms constitute well-rehearsed routines that are found across many infants' data. A longitudinal perspective will address issues that remain from the previous chapter regarding how the infants' language ability shapes these interactions over time, as language faculties develop to allow infants to play a vocal role in these turn-taking dialogues. In line with this, we will also observe the caregiver's OW and CW production, noting any changes that take place in the infants' input as output capacities develop. The concept of wildness will also be reconsidered, in order to determine whether this is a true feature of OW production, as proposed in Rhodes' (1994) account; this will return us to the questions raised in Chapter 4 regarding the role of wildness in determining iconicity, and whether or not this may facilitate infants' perception of OWs. This broad analysis of the longitudinal and contextual scope of onomatopoeia allows us to consolidate the findings observed in this thesis, addressing some of our unanswered questions in the process.

Eight infants acquiring either American English or French will be analysed in order to determine the nature of OW production in early interactions from a longitudinal perspective. First, an overview of the use of OWs and CWs in relation to the infants' general language development will be discussed in comparison with the caregivers' use of OWs (Study 1). Then, OW production will be considered alongside the use of external stimuli such as books and toys (Study 2), followed by a close-up analysis of OW and CW production in turn-taking interactions (Study 3). Finally, in Study 4 we will return to phonological wildness in order to explore the reality of this feature in caregivers' OW production, and how this might facilitate production in the early infant output. Together these varied analyses will paint a realistic picture of OW use from both an input and an output perspective, supporting the analyses that have already been carried out in this which show a dynamic and experience-based role for OWs in early language development.

7.2 General methodology

7.2.1 Participants

Data were collected from longitudinal corpus studies in the Child Language Data Exchange System (CHILDES, MacWhinney, 2000) database, and were selected based on the availability of data from the onset of word production until the use of word combinations. Details of the infants analysed and the corpora from which the data was sourced are shown in Table 7.1.

Table 7.1: Infants analysed in the present study.

Corpus	Infant	Gender	Language
Providence -	Lily	F	US English
(Demuth, Culbertson & Alter, 2006)	Naima	F	US English
	Alex	M	US English
	William	M	US English
Lyon - (Demuth & Tremblay, 2008)	Anais	F	French
	Marie	F	French
	Nathan	M	French
Paris - (Morgenstern & Parisse, 2007)	Théophile	M	French

Infants were all typically-developing and acquiring English or French as their only language.

Three recording sessions were analysed for each infant; in all but one of the recordings the infants were interacting with the mother as the main caregiver, with occasional input from the researcher, father or siblings. Théophile is the only infant whose father was consistently present during the three sessions: in Sessions 1 and 3 the mother has the most prominence in the recordings and features in almost all adult utterances and caregiver-infant interactions, but due to the mother's illness the father takes the lead in Session 2. Rather than discount this recording from the dataset it was decided to include it in the analysis; this again provides an accurate picture of the reality of the infants' early language experience, which in most cases includes extensive input from adults other than the mother. However, interactions involving researchers, grandparents, siblings and any other interlocutors featured in the recordings were not analysed.

7.2.2 Recording procedure

Each infant was recorded by the original experimenter in the home for one hour on a monthly (Paris corpus) or fortnightly (Lyon and Providence corpora[15]) basis from the onset of first words. A video camera was set up in the home to record interactions between the infants and their caregivers, and in most cases an experimenter was present

[15] Naima (Providence corpus) was recorded weekly between the ages of 1;3 and 2;10.

to take the recordings. The mothers and their infants wore microphones (in some cases the infant wore the microphone in a small backpack) to provide clear audio data. Recordings and full transcriptions were made available on CHILDES, including detailed phonetic transcriptions of the infants' utterances.

7.3 General analysis

Three recording sessions were analysed for each infant. Session 1 aimed to capture the earliest stage in the infants' word production, and so the first session in which the infant produced at least five different *word types* was analysed (mean word types = 14.13). William is identified as an anomaly in the first session, as although he fits within the group in terms of age, his data collection is reported to have started late (Demuth et al., 2006), and this is reflected in his high word production here (see Table 7.2). This will be considered in the analysis of William's early data.

Session 3 was then established as the first recording in which the infant produced at least five *word combinations*. A phrase was considered to be a word combination if at least two words were produced spontaneously to constitute a meaningful utterance, and if this utterance was not repeated from or prompted by the caregiver. The five word combinations also had to be distinctly different: *yellow duck* and *yellow car* would not be considered as distinctly different word combinations, for example, and so would instead be counted as one. Likewise, *where's the sheep?* and *where's the cow?* would be considered as one example of a word combination in terms of the lexical items that they consist of. Furthermore, 'frozen' phrases such as *what's that?* are considered as single lexical units, and do not count as examples of word combinations.

Finally, Session 2 was determined by identifying the mid-point in each infant's age between Session 1 and Session 3. This provides an intermediate developmental stage between the onset of productive word-use and the beginning of word combinations. The infants' ages and extent of their word production in each of the three sessions can be seen in Table 7.2.

Table 7.2: Age and word production in each session across infants.

	Session 1		Session 2		Session 3	
Infant	Age	Word types	Age	Word types	Age	Word types
Lily	1;4.28	6	1;7.7	14	1;9.8	31
Naima	1;1.2	11	1;5.8	8	1;8.14	104
Alex	0;11.27	14	1;1.11	20	1;3.12	137
William	1;4.12	53	1;7.18	62	1;9.12	101
Anais	1;2.11	6	1;5.19	27	1;8.11	19
Marie	1;0.2	9	1;5.10	20	1;7.12	41
Nathan	1;2.26	6	1;7.8	11	1;11.3	24
Théophile	1;5.5	8	1;9.6	14	2;0.20	33

As Table 7.2 shows, there is wide variability in both the infants' ages and the number of word types produced by each infant across the three sessions. This provides an opportunity to determine how OWs are used by the infant and the caregiver across more and less voluble infants in the early stages of language production.

Both phonetic- and utterance-level information provided by the original transcribers has been adhered to in this analysis. All transcriptions were cross-checked with the video recording by the current author to ensure that the data was accurate, and any amendments were made accordingly. In some cases an OW was not recognisable from the transcription alone (for example, the transcription [m̩ m̩ m̩] recorded as *choo choo* in William's data (Session 1)) and in all such cases this was confirmed from an impressionistic phonetic analysis of the recordings. In cases where multiple utterances occurred from one speaker (usually the caregiver) without interaction from the other, pauses can typically be found which demarcate one utterance from the next. OWs are the main consideration in this analysis, but as has been the case in the previous chapters, this will be compared with the use of CW forms where possible. However, not all instances of CW forms will be extracted from each recording here; instead, only those CWs which occur in the presence of one or multiple OWs will be considered, as CWs may be produced elsewhere in the recordings in situations unrelated to OW production. Furthermore, unlike the previous analyses, OW-CW pairings will not be considered; instead the use of OW forms in general will be analysed, with a consideration of how CWs fit within these interactions to broaden our understanding of OW and CW use in early language development. Naturally we must consider that in this case the analyses of CW production are not necessarily representative of the use of these forms overall, but here they provide evidence towards our understanding of OW acquisition as part of the wider context of general language development.

Owing to positive skewness across the dataset, which led to non-normal distributions in both the caregivers' and the infants' data, a log10 transformation was applied to all results. Figures represent the original non-transformed data.

7.4 Study 1 – OW production from a longitudinal perspective

Study 1 seeks to understand the relationship between the infants' and caregivers' production of OWs in early interactions. This will function as an extension of the analyses carried out in Chapters 2 and 3, with the additional perspective of the caregiver's role in an infant's acquisition and use of OWs. Observations of the caregiver's OW production alongside their infants' developing lexicon will allow a more in-depth analysis of the changes in OW production which take place over the course of the second year of life. This will also provide us with a broader view of the changing nature of OWs in the early input as well as the infants' output.

Based on evidence from the previous chapters, we hypothesise here that the input plays an integral role in the infant's developing output. This analysis will expand on results from Chapter 6, to identify the frequency and variability of OWs in the caregivers' speech. In line with Kauschke and colleagues' findings (2002, 2007), here we predict that exposure to OWs in the input will prompt OW production in the infant output, and that this may also be determined by factors affecting the caregivers' OW production – infant age and language ability, for example. Indeed, we expect the nature of the caregivers' OW use to change over the course of the three sessions, determined by the infants and their increasingly independent and voluble word production.

7.4.1 Analysis

Each dyad was analysed for the total number of OWs and corresponding CWs produced in each session, alongside overall word production. Conventions from Chapter 6 relating to the use of reduplication were adhered to, and so reduplicated OWs were considered as one token, while pauses between identical words (as transcribed by the original researchers) reflected repetition of multiple tokens. Forms which were read from a book, and so pre-determined in the caregivers' output according to a 'script', were not considered in this analysis.

7.4.2 Results

OW use and language ability: caregivers

OWs were produced to some extent by all eight caregivers in each session. This confirms the reality of OW use in IDS, although variability was observed across the dyads. A linear mixed-effects model was generated in R with caregivers' OW production as a dependent variable, infant age and sex as fixed effects and dyad as a random effect. By-subject random slopes were included for the effect of age. No effect was found on the number of OW tokens produced by the caregiver for either infant age in months ($\chi2$ (1) =.000, p=.99) or sex ($\chi2$ (1) = 1.21, p=.27), and so these factors were not included in further analyses. The same test was carried out with session as a fixed effect and by-subject random slopes for the effect of session. No effect was found for session on the caregivers' OW production ($\chi2$ (1) = .297, p=.59), suggesting that the infants' developing language ability was not a factor in the number of OW tokens produced by each caregiver in each session. The distribution of OW-production across each dyad is shown in Figure 7.1.

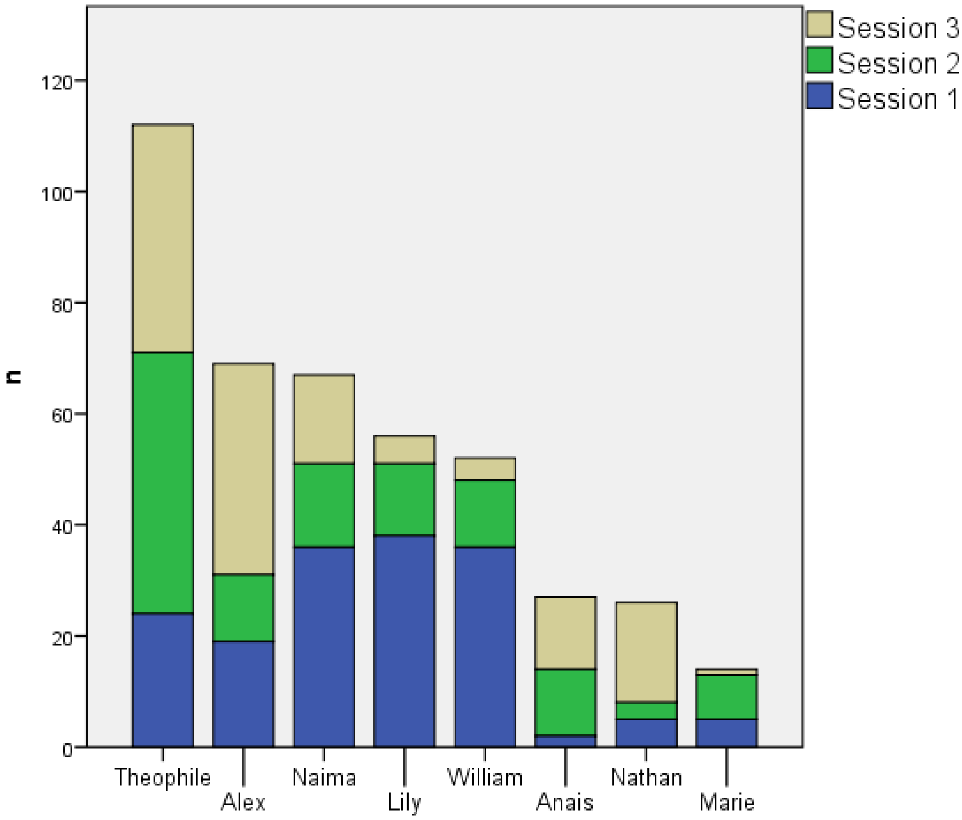

Figure 7.1: OW tokens produced by caregivers across the three sessions.

Pearson product-moment tests revealed significant positive correlations between the caregivers' production of OW types and OW tokens across all sessions (Session 1: $r = .883$, $n = 8$, $p = .004$; Session 2: $r = .772$, $n = 8$, $p = .025$; Session 3: $r = .957$, $n = 8$, $p < .000$); that is, those who produced a wider variety of OWs also produced a higher volume of these forms.

We can also see from Figure 7.1 that the three infants whose caregivers produced the lowest number of OWs are all French. However, while this could be assumed to point to language-specific trends in OW production, Théophile's caregivers produce the highest number of OWs across the entire dataset. Effect of language on the caregivers' OW production was tested using a linear mixed-effects analysis. The caregivers' OW production (tokens) was included as the dependent variable, with fixed effects of language (French or English) and session, and dyad as a random effect, including by-subject random slopes for session. No significant effect was found for language in the caregivers' production of OW tokens across sessions ($\chi 2$ (1) = 1.71, $p = .19$). However, when the same model was generated without Théophile's data, a highly significant effect was found for language on the mothers' OW production ($\chi 2$ (1) = 12.97, $p < .000$); on average, American mothers produced 14.8 more OW tokens per session than the French mothers.

OW use and language ability: infants

The infants' OW production was then analysed across sessions. Linear mixed-effects models compared the infants' overall word production in each session. Age and sex were included as fixed effects and dyad as a random effect, with by-subject random slopes for the effect of age. This model revealed a strong effect of age on the infants' outputs, both in terms of overall lexicon (total number of word types: $\chi 2$ (1) = 13.795, $p < .000$) and total number of OWs produced (OW tokens: $\chi 2$ (1) = 4.936, $p = .026$). No effect was found for sex on either overall word production ($\chi 2$ (1) = 0.195, $p = .66$) or OW production ($\chi 2$ (1) = 0.007, $p = .93$), and so this was not considered as a factor in further analyses. Language was then considered as a factor: a linear mixed-effects model was generated with total word tokens as a dependent variable, language (English and French) and session as fixed effects, dyad as a random effect and by-subject random slopes for the effect of session. Language was found to have an effect on the infants' overall word production ($\chi 2$ (1) = 4.046, $p = .044$): on average French infants produced 29.3 fewer words than the American infants across sessions. Finally, consistent with the analysis of the caregivers' use of OWs, language had no effect on the infants' OW production when all infants were included in the model (χ2 (1) = 1.634, p=.2), but when Théophile's data was removed from the analysis, a strong effect was found (χ2 (1) = 7.95, p=.004); on average, American infants produced nine more OW tokens than the French infants per session.

The infants' production of OWs across the three sessions was then analysed. A linear mixed-effects model compared infants' OW production (types), with the fixed effect of session and dyad as a random effect, with a by-subject random slope for session. A significant effect was found for session (χ2 (1) = 4.675, p=.031) on the infants' OW production: as expected, infants produced an increasing number of OW types over the course of the three sessions. However, paired-samples t-tests showed this difference to occur between Sessions 2 and 3 (t(6) = -3.404, p=.014), while no difference was found between the infants' production of OW types between Sessions 1 and 2 (t(6) = .125, p=.905). This is illustrated in Figure 7.2.

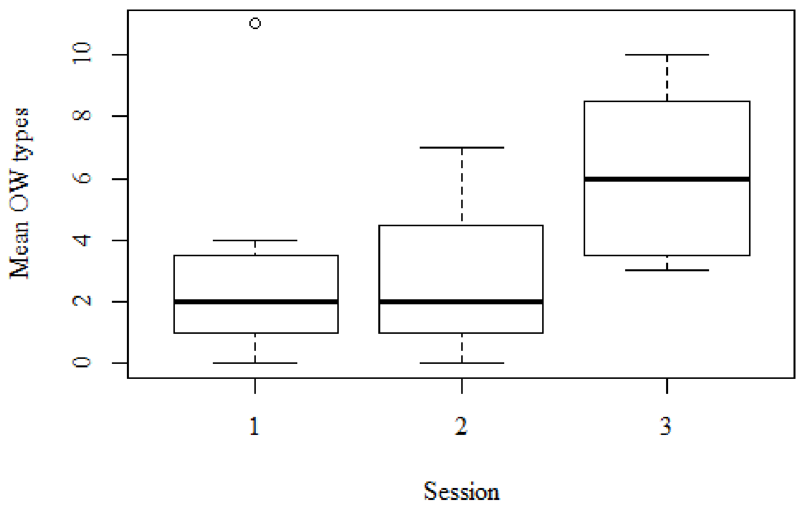

Figure 7.2: OW types produced by the infants across sessions.

When number of OW types was calculated as a proportion of each infant's overall lexicon, the same mixed-effects model as generated above but with proportion of OWs as a dependent variable revealed no difference between infants' production of OWs over the course of the three sessions (χ2 (1) = 1.91, p=.17). This corresponds to findings from Chapter 2 showing a steady overall decrease in the acquisition of OWs over the course of early production: OW acquisition was most prolific before the 40-word point, and then tailed off as the infants acquired an increasing number of 'regular' words (RWs). As Figure 7.3 shows, a similar trend can be observed in this data, with a subtle decrease in the proportion of OW types produced over the three sessions (Mean Session 1: 19.12% OWs, Session 2: 15.17% OWs, Session 3: 14.88% OWs). Again, this is consistent with results from Chapter 2, where we saw how OWs accounted for on average around one third of the infants' earliest words before decreasing to less than 10% by the 100-word point.

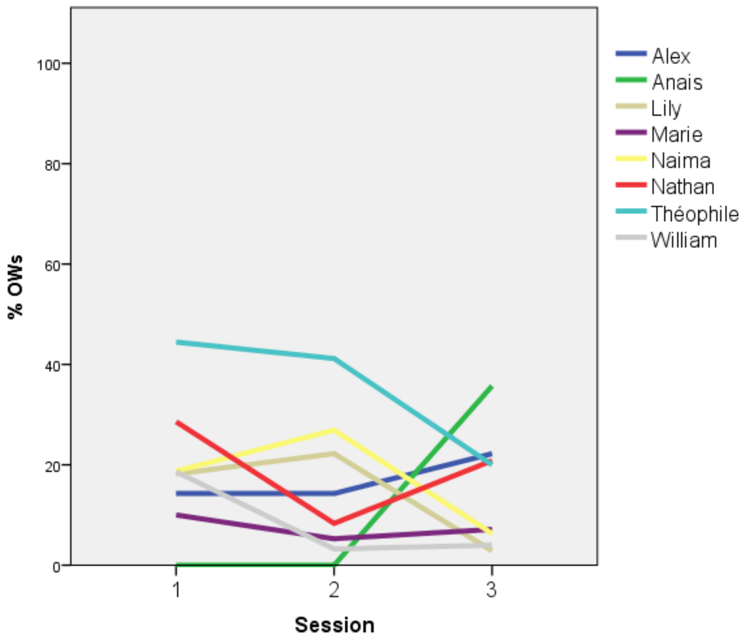

Figure 7.3: OW types as a percentage of all word types produced in each session.

Individual differences are clear from Figure 7.3 – Anais in particular shows the opposite trend to what is expected here – but overall we see a downwards trajectory in the extent of the infants' OW production. As discussed in Chapter 2, this is to be expected; once lexical development begins to increase more rapidly, OWs have a smaller influence on the overall output as the infant has a wider range of lexical items to draw upon in interactions with the caregiver.

OW production across infant and caregiver

The caregivers' OW production has been shown to be independent of their infants' language ability. This was independent of the individual infants' lexical capacity in any one session, as well as the general shift forwards that we expect to see in all infants' language development over the course of the three sessions. Similarly, no consistent trends have been found across the infants' OW production and their own language development. However, when comparisons are made between the infants' and the caregivers' production of OW tokens in Session 1, a direct relationship can be identified between the two: a Pearson product-moment test reveals a significant correlation between the number of OW tokens produced by the caregivers and their infants in this session ($r = .801$, $n= 7$, $p=.03$); those infants whose caregivers produce few OWs in the input also tend to produce a small number of OWs, and this is consistent with those caregivers who produce many OWs (Figure 7.4). The same trend is also observed across Sessions 2 and 3 (Session 2: $r = .776$, $n= 7$, $p=.04$; Session 3: $r = .756$, $n= 8$, $p=.03$).

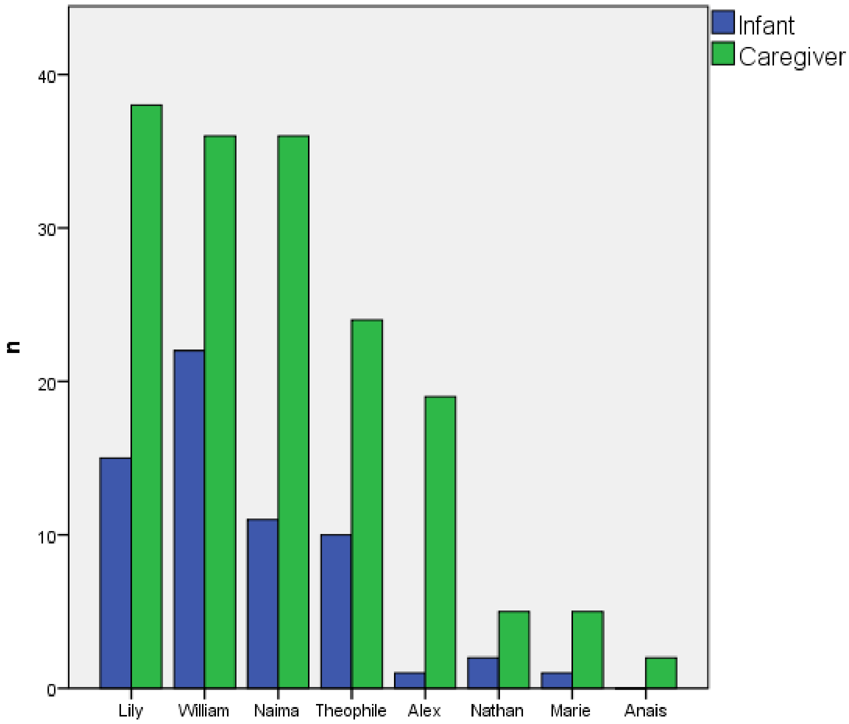

Figure 7.4: Number of OW tokens produced in Session 1.

Similarly, the variety of OWs produced by the mothers in Session 1 differs across dyads: William, Alex and Lily were all exposed to between 13 and 16 different OW types in this session, whereas Marie and Anais heard only three and two OW types, respectively. While the production of many OW tokens in the infant input has an effect on OW-use in the early output, the range of different OW types has less of an influence: no correlation can be found between the number of OW types produced by the caregivers and their infants in Session 1 ($r = .425$, $n= 7$, $p=.342$). However, the same test shows a significant correlation in the production of OW types in Session 2 ($r = .776$, $n= 7$, $p=.04$) and Session 3 ($r = .874$, $n= 8$, $p=.005$), showing that variability in the input becomes an increasingly important factor determining the infant's output over time.

Finally, a between-sessions analysis showed no connection between the caregivers' OW use in the early stages of word production and the infants' later use of OWs. A Pearson product-moment analysis shows no correlation between the number of OW types or tokens produced by the caregivers in Session 1 and the number of OW types or tokens produced by the infants in the following sessions (types - Session 2: r = .267, n= 7, p=.536; Session 3: r = -.202, n= 8, p=.631; tokens - Session 2: r = .496, n= 7, p=.257; Session 3: r = .179, n= 8, p=.671).

OWs and CWs in the infant input

In Chapter 6 we reported that OWs are rarely presented to the infant without a corresponding CW to 'support' the OW. These results are confirmed across the eight dyads analysed here: a significant correlation can be found between the caregivers' production of OWs and CWs in terms of both word types (r = .75, n= 22, p=.000) and word tokens (r = .554, n= 22, p=.007); across all three sessions the caregivers are found to produce CWs in tandem with their corresponding OWs. Furthermore, when Théophile's data is excluded from the analysis we see even stronger correlations in both conditions (word types: r = .880, n= 19, p=.000; word tokens: r = .729, n= 19, p=.000): as Figure 7.1 shows, his caregivers (both mother and father) produce a particularly high number of OWs, and often these are not accompanied by a CW. The nature of Théophile's OW production is unique in this dataset, and will be considered in more detail in the analyses that follow.

Whereas correlations have been found between the caregivers' and infants' OW production across all three sessions, the same trends cannot be found in infants' CW production. Owing to a lack of CWs in the infants' data, comparison within the dyads is not possible here; this again supports results from Chapters 2 and 3 which show that infants acquire OWs early on in their development but produce the equivalent CWs only later. However, a comparison of the caregivers' CW production against the infants' OW production over time reveals some interesting results. While there is no correlation between the caregivers' CW types and the infants' OW types in Session 1 (r = .166, n= 7, p=.722) or 2 (r = .668, n= 6, p =.147), a correlation can be found in Session 3 (r = .833, n= 8, p=.01). As the infants' OW production increases, caregivers produce an increasing number of CWs. This is shown in Figure 7.5, where we see how the infants' OW production eventually exceeds that of the caregivers.

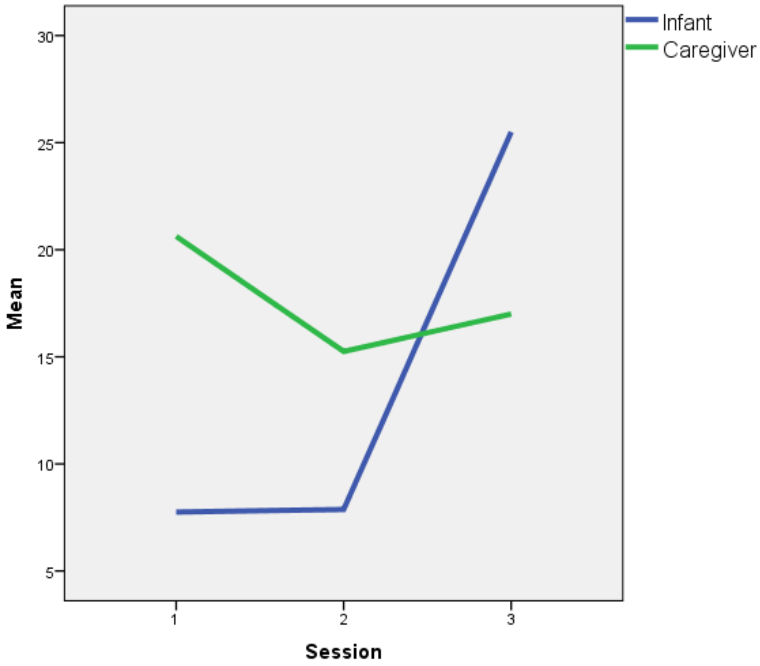

Figure 7.5: Interaction between infants' and caregivers' production of OW tokens.

Two two-way repeated-measures ANOVAs were carried out in R with OW types/tokens as dependent variables to test the interaction between infants' and caregivers' production across sessions. Speaker (infant vs. caregiver) and session were included as factors. Significant interactions were found between the infants' and caregivers' OW production across sessions in both cases (types: $F(1, 6) = 8.406$, $p=.007$; tokens: $F(1, 6) = 10.567$, $p=.003$), suggesting that caregivers are moving away from OWs and towards more adult-like CW production just as the infants begin producing OWs with more consistency.

7.4.3 Discussion

In this analysis we have observed both variability and consistency across the eight dyads' OW production. Perhaps the most important finding to note is that OWs were produced across the board, with each caregiver producing at least some OWs in each of the three sessions. However, some were found to produce a high output of OWs across sessions while others consistently produced very few OWs in interactions with their infants. Furthermore, caregivers who produced a high quantity of OWs were also found to use a wide variety of these forms: while some infants receive a rich input from their caregivers with both a high quantity and a wide range of OW types, the opposite is true for others. It is clear that OW use is specific to the individual dyads, and this goes on to influence each infant's eventual OW production.

Correlations were found between the caregivers' and infants' outputs across all three sessions. In Session 1 a strong connection was made between the number of OWs presented to the infants in the input and the extent to which they produced OWs in the same session. At this stage the infants were at the very beginnings of word production, and so we see how integral the caregiver's speech is to the newly-vocal child. While the quantity of OWs heard in the input is the determining factor in Session 1, in Sessions 2 and 3 variability becomes a more important feature of the caregivers' production, as we see strong correspondences between the number of different OWs produced within the individual dyads. These results reflect similar findings from Kauschke and colleagues (2002, 2007), who also observed parallels between caregivers' and infants' OW production over time.

It could be suggested that the caregivers match their word production to the changing abilities of their infants, which may explain the interaction identified in infants' and caregivers' output forms over the course of the three sessions. As the infants' lexicons become more varied, so too might the caregivers' outputs increase in variability, thus prompting the production of a wider range of OW forms in the later sessions. However, it was shown that the caregivers begin to move away from OW production as the infants' lexicons become increasingly sophisticated, and their use of CW forms is found to match their infants' production of OWs. It could be suggested that we are observing evidence for the 'zone of proximal development' here (Taumoepeau & Ruffman, 2008; Vygotsky, 1978; Zimmerman et al., 2009), as caregivers interact dynamically with their infants' changing abilities. Evidence for this was also found in the infants' output, as we continued to observe their production of OWs rather than the CWs of their caregivers, suggesting that the caregivers' speech is "just challenging enough" (Zimmerman et al., 2009, p.347) to promote language development, without going too far beyond the infants' current language abilities. Indeed, the caregivers do not stop producing OWs in these later sessions, as we still see correlations over time, which relate OW as well as CW production in the input to the infants' use of OWs in the output: the variability observed in the caregivers' OW production was found to determine the variability of the infants' OW production. These findings bring us back to the nature of interactions and turn-taking routines in early language development, raising the question of how OWs and CWs are produced in infant-caregiver routines. This will be explored further in Study 3.

Results from this study can be related to observations from previous analyses in this thesis. Similarities were found with regard to infants' early production of OWs in relation to their wider vocabulary and the changing nature of the lexicon over the course of development. As in Chapters 2 and 3, OWs were initially a dominant feature of the early vocabulary, accounting for around 20% of the output in Session 1. This decreased over the course of the analysis, and OWs accounted for around only 14% of the infants' output forms by Session 3. While these results are not as convincing as the sharp longitudinal decrease in OW production observed in Figure 2.2, the range of data is comparable nonetheless, with OWs accounting for up to 44% of the output in Session 1 (observed in Théophile's data) to as little as 4% in Session 3 (William's data): again, the nature of OW use is specific to the individual infant, and we observe the same variability across the infants in Chapter 2. Additionally, we were unable to compare the caregivers' CW production with that of their infants, as, despite them all producing a range of OWs, very few of the infants produced any CWs during the video recordings. This also supports findings from Chapters 2 and 3, showing how the early acquisition of OWs appears to delay the need for CW production. Of course, without a full account of each infant's lexicon we are only able to speculate as to their overall vocabulary development based on the data provided in the recordings.

From this analysis we can observe the dynamic effect of the caregiver input on the infant output, as caregivers adapt their speech over time to match the infant's abilities while also prompting further linguistic development. We can also observe how OW production changes over time, in both infant and caregiver speech, as the infants' overall language ability becomes more sophisticated. The correspondences identified within sessions, paired with the lack of correspondences across sessions, highlights the time-locked nature of these interactions, as the ever-changing infant output adapts on-line, first to the quantity and then to the variability of the language presented in the input.

7.5 Study 2 – Context of OW production

Chapter 6 presented data from mothers reading to their infants from two selected picture books, where we found that images of animals, vehicles and other objects with potential for OW-use prompted the mothers' OW production in the book-reading task. However, this does not give us a holistic view of the use of OWs in the home, as this predetermined activity provides only a narrow perspective on the use of OWs in interactions. This leads us to question the reality of OW production, and whether it is exclusive to book-reading activities, or whether it occurs as part of the infants' general input across varying contexts. Yont and colleagues (2003) observed that OWs were produced by infants during toy-play activities, and the authors propose that these forms may be used by the infants as a 'conversational tool' (p.448) in the absence of established conversational competence. Indeed, it is suggested that toy-play, as opposed to book-reading, offers infants more opportunities to produce a spectrum of lexical and syntactic features. This raises the question of how the context of OW use may promote language learning and, importantly for this analysis, how it defines OW production.

In this study we will analyse the use of OWs and CWs in infant-caregiver interactions more closely, in order to determine the contexts in which OWs are delivered to infants in the input. Here we will consider whether OW production is prompted by the use of external stimuli such as books and toys, or whether these forms are independently produced in vocal interactions without the need for any contextual prompting. This will enable us to understand how OWs are used in everyday caregiver-infant interactions, either as context-bound forms which are specific to particular activities and stimuli, and thus limited in the input, or as a more general feature of early word production. From this analysis we will be able to establish whether OWs are always produced with an accompanying referent (e.g. *woof woof* occurring only in the presence of a 'dog' referent) – potentially a highly relevant aspect in the development of form-meaning correspondences throughout lexical acquisition – and how common these forms are in activities of joint attention.

7.5.1 Analysis

OW and CW exchanges

Independent utterances were coded for the use of OWs and CWs. The coding system used was an adaptation of Ninio and colleagues' (1994) INCA-A (Inventory of

Communicative Acts-Abridged) system, which codes communicative intent at two different levels – the verbal level, representing the general context of an utterance (or 'interchange'), and the utterance level, representing its speech-specific features. For the purposes of this study, these 'interchanges' are termed 'exchanges', and Ninio et al.'s two levels are considered together as a homogenous group of features, which will be identified individually in this analysis and in the two studies that follow. In line with the INCA-A system, an exchange is defined here as an utterance which serves as a "unitary interactive function" (Yont et al., 2003, p.440), involving either the *production* (spontaneous or imitated) or the *prompting* (explicit or incidental) of an OW or CW by either the caregiver or the infant. That is, an exchange is a single utterance that forms part of a wider interaction. Utterances in which the caregiver asks *what does the cow say?* or *what's that?* while pointing to a toy or image of a cow would both be considered as CW exchanges: in both cases the focus of the phrase is the CW *cow*, while the expected response from the infant would be the OW exchange *moo* in the first case and the CW exchange *cow* in the latter. OW exchanges involve either the production of or focus on an OW, such as *he says woof woof* or *who says moo?* Importantly, it is not necessary for a phrase to contain an OW or CW to be considered as an OW/CW exchange, but the phrase must refer unambiguously to an OW/CW through questioning if it is not produced outright. This allows for an analysis of the nature of dialogues in which OW and CW forms occur: exchanges do not always feature the production of the form in question, and so utterances which attempt to prompt OW or CW production must also be considered, as these exchanges are also highly relevant to this analysis.

Only 26 exchanges in the dataset contain both an OW and the corresponding CW, as in *lion you have the loudest roar* (Lily, Session 1). These are rare owing to the use of pauses in OW production, as observed in Chapter 6: while many OW exchanges are followed by a corresponding CW exchange, we often find a pause after an OW, which leads to a natural division between utterances. Consequently, these are considered here as separate exchanges[16]. We see an example of this in (2) below, where an OW exchange is followed by a pause, with a CW exchange immediately afterwards:

(2) Session 1: Naima

```
MOT:   buck buck bagah (.)
MOT:   that's what a chicken says
```

[16] They are also recorded as separate utterances in the original transcripts.

The 26 combined OW-CW exchanges are termed combination exchanges in this analysis, and the OW and the CW portion of the exchange will each be considered individually, as discussed in further detail below. Combination exchanges appear exclusively in the caregivers' data.

Despite the fact that the caregivers have already been found to produce similar numbers of OWs and CWs, we see more OW exchanges than CW exchanges in this dataset: across all three sessions a total of 1,133 exchanges are recorded (710 OW, 397 CW), as well as the 26 combination exchanges. However, it should be reemphasized that the number of exchanges does not reflect the total number of OWs and CWs produced in the dataset, as was previously analysed in Study 1. Many exchanges constitute multiple instances of the same word form, which in this analysis are considered as one part of a wider interaction between infant and caregiver and are thus counted as a single exchange. In contrast, many exchanges show the caregiver attempting to elicit production from the infant without producing the OW or CW explicitly. Figure 7.6 shows the total number of OW and CW exchanges produced by the caregivers and their infants across all three sessions.

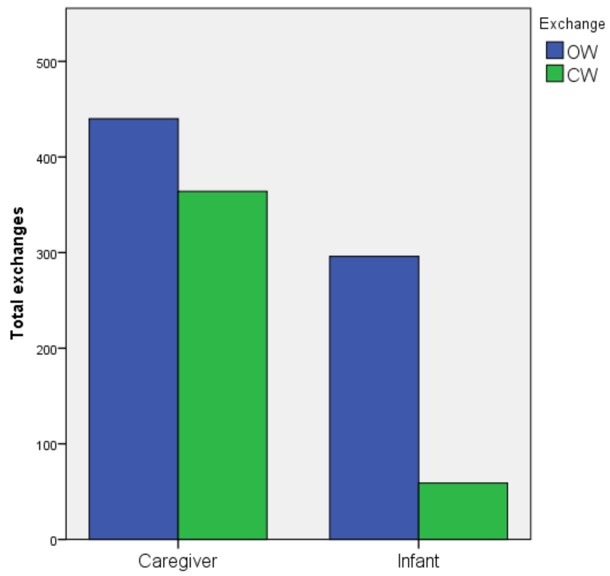

Figure 7.6: OW and CW exchanges produced across sessions.

External stimuli

The dataset was coded for the use of external stimuli during an exchange, which enabled us to determine the context of OW and CW production. External stimuli are detailed in Table 7.3.

Table 7.3: Definition of external stimuli used.

Stimuli	Situation
IMAGE	reading books, colouring books, looking at photos
OBJECT	playing with toys, balloons, swings and other playground items
ACTION	tickling, eating, jumping, dropping an object
ROUTINE	singing, games such as peek-a-boo
NONE	no identifiable context

Exchanges taking place in situations where the use of stimuli was unclear because the infant was either out of view of the camera or in a position where the stimuli couldn't be seen were not counted in this analysis. One section of data from Théophile's recording in Session 3 is not included here due to an extensive period of crying, although a large number of OWs were produced by both caregiver and infant during this section.

7.5.2 Results

External stimuli and OW production

OW exchanges were accompanied by an external stimulus (IMAGE, OBJECT or ACTION) in 97% of cases, and six of the eight infants' exchanges involve an external stimulus in 100% of cases across all three sessions. For all infants, IMAGES and OBJECTS are the most common stimuli used, accounting for 41% and 39% of all exchanges in the dataset, respectively. Exchanges involving no external stimuli (NONE) account for 9% of the full dataset, but these are exclusive to Naima and Théophile's data, occurring in all of Naima's sessions and two of Théophile's sessions. Individually, these exchanges constitute 46% and 11% of these infants' total exchanges, respectively. Only William's mother makes use of ROUTINES in the production of OW exchanges (always through singing). ACTIONS only occur in three infants' data (Marie, Nathan and Théophile), and account for almost a third (32%) of Théophile's OW exchanges. Here we see variability across dyads even in terms of the context in which OWs are produced. The use of these different stimuli will be discussed in more detail below[17].

Objects

Forty five percent of all OW exchanges are produced during OBJECT-based activities such as toy-play. When the data is considered more closely it is clear that OWs are

[17] Routines is omitted from further discussion as these exchanges occur only in the production of songs.

produced in a specific manner during toy-play activities. These forms are used by both infants and caregivers as an accompaniment to the infants' play, and so do not constitute the main focus of the situation at hand, which is in all cases supported by the toy. Because of this, the production of the accompanying CW is often unnecessary, as the main interaction takes place between the infant/caregiver and the OBJECT in question. Examples from the dataset can be seen below.

(3) Session 1: Naima

```
MOT:   and here's mister woof who's saying +...
MOT:   woof woof woof woof woof woof woof.
[SIT:   MOT moves toy towards CHI as she speaks]
```

(4) Session 2: Théophile

```
MOT:   ah tu veux que je te pousse.
        ah you want me to push you.
MOT:   vroum vroum vroum.
        vroom vroom vroom.
[SIT:   pushing CHI in toy car]
```

So far in this thesis we have observed caregivers' consistent production of CWs to accompany OWs, but these examples show that this is not always the case. Here we see the extent to which OW exchanges can take place independently of CW production, as the toy is used in place of the CW to provide context, thus 'supporting' OW production here. After the caregiver's initial OW exchange in example (3) she goes on to produce a further eight OW exchanges, with no corresponding CW exchange. However, we do see labelling where a CW could otherwise be presented, in the use of *Mister Woof* in reference to a toy dog. The same is true of example (4), which continues with nine further OW exchanges produced by both the caregiver and the infant, all in reference to the toy car which is never explicitly mentioned – the physical presence of the OBJECT alongside the OW exchanges provides ample context for this interaction.

As these examples show, the OWs are implemented as part of a play situation, and not as an aspect of vocal interaction between caregiver and infant. The caregivers produce OW exchanges as a commentary to the infants' toy-play, but the physical OBJECT's role in the interaction leaves no ambiguity and thus replaces the CW here.

Images

In total, interactions involving IMAGES account for 31% of all OW exchanges, and 56%

of all CW exchanges. Examples (5) to (7) demonstrate the typical exchanges that take place during book-reading activities:

(5) Session 1: Lily

```
MOT:   the lion.
MOT:   what does a lion say?
CHI:   arr [=roar]
MOT:   a sheep.
MOT:   baa baa:.
MOT:   a mouse!
MOT:   squeak squeak.
MOT:   a pig.
MOT:   oink oink.
```

(6) Session 2: William

```
MOT:   what's, what's that?
CHI:   dogs.
MOT:   that's a dog, that's right, what's a dog say?
CHI:   woof woof.
MOT:   woof woof.
MOT:   that's right William
MOT:   that's a cow, what's the cow say?
CHI:   cow.
MOT:   uh huh, what does a cow say?
MOT:   cow.
MOT:   what's the cow say William?
```

(7) Session 3: Alex

```
MOT:   those are the ducks.
MOT:   what does the duck say?
CHI:   duck.
MOT:   what does the duck say?
CHI:   quack quack.
MOT:   yay, quack quack quack.
```

The first point to be noted is the striking similarity across these three examples, which contain similar interactional features as those observed in Chapter 6 (examples (2) to (4)). The caregivers make extensive use of questioning here in an attempt to elicit word production (both OW and CW) from their infants and engage them in turn-taking dialogues. The book-reading activity serves to provide visual prompts for word production between the caregiver and the infant, but in this case both CW and OW forms are used alongside the external stimulus.

Examples (6) and (7) demonstrate multiple attempts from the mothers to elicit OW production specifically: even though, through repetition, both infants produce the correct CW form in both instances, the mothers continue to pursue the OW exchange. This suggests that the mothers may be pushing the infants forwards in their word production at this stage, moving them away from repetition (and reliance on the immediate input from the mother) and towards a more independent role in interactions.

Actions

Examples of exchanges involving ACTIONS are more limited in the dataset as they tend to be spontaneous and therefore one-off events. In the majority of cases (81%) these are OW exchanges; ACTION exchanges account for 17% of all OWs in the dataset, and 6% of all CWs. In many cases no real interaction takes place between the infant and the caregiver; we typically find the exchange to be peripheral to the situation at hand. These points are demonstrated in examples (8) to (10) below.

(8) Session 1: Nathan

```
CHI:   boum.
SIT:    [CHI sits on floor]
```

(9) Session 2: Théophile

```
CHI:    pum.
       boum
SIT:   [kicking flowerpot]
```

(10) Session 3: Anais

```
CHI:    vroum. [=car sound]
SIT:   [running around room]
```

Ninety six percent of all ACTION exchanges (n= 146) are produced in Naima and Théophile's recordings: OW exchanges involving ACTIONS account for around one third of each of these infant's total exchanges (both OW and CW). This highlights the variability across dyads in terms of individual approaches to OW production, as some infants produce no OWs in relation to ACTIONS, while others can be found to rely heavily on this context.

No external stimuli

As reported above, Naima and Théophile are the only dyads to produce exchanges in the absence of an external stimulus. There are 22 such instances in the dataset, and 77% of these exchanges occur in Théophile's data. Here we find impromptu exchanges taking place as responses to one-off events in the infants' surroundings. In both cases we see examples of incidental vocal production, whereby the infants respond to their environment spontaneously with words they can produce (both OWs and CWs) without any prompting from the caregiver, and with no joint attention routine involving an IMAGE or OBJECT stimulus. For example, Théophile responds to a banging sound in the background of the recording with a spontaneous production of *boum* 'boom', to which the caregiver responds *oui tu as raison ça fait boum* 'yes you're right, it goes boom'. Similarly, Naima responds to her mother's comments about it being *seven o'clock, time for bed* with both a CW and an OW exchange: *clock…tick tock tick tock tick tock*, but in this case her mother does not attempt to repeat these vocalisations, instead confirming that it's *time for bed*.

7.5.3 Discussion

The main finding to be observed in this analysis is that OW exchanges occur in the presence of an external stimulus in almost all cases (97%). This points to a heavily context-bound role for these forms which, more often than not, involves a physical depiction of the referent in question. The use of OBJECTS most often accompanied OW production for both infants and caregivers, followed by IMAGES; this is in accordance with the literature, which shows both book-reading and toy-play situations to provide the most variable linguistic input, in terms of both infant and caregiver speech (Soderstrom & Wittebolle, 2013; but see also Rondal, 1980; Yont et al., 2003). However, the nature of these exchanges differed according to those external stimuli: book-reading activities included vocal interaction between the infant and caregiver, with conventionalised turn-taking routines that prompted infant word production. On the other hand, the production of OW/CW forms during toy-play was largely supplementary to the play situation, and often involved making appropriate sound effects rather than constituting any turn-taking dialogues. Importantly for both of these contexts, the physical use of a toy or attention to a picture in a book was consistently combined with OW/CW production, and so we can propose that the interaction provides an increased learning opportunity owing to the presence of the physical referent.

In the book-reading interactions we observed further examples of the routinized use of CW and OW exchanges, as initially noted in Chapter 6. This is consistent with Yont and colleagues' (2003) comments regarding OWs in early interactions, which were found to support infants' participation in turn-taking dialogues. However, we also observed examples of infant CW production when the caregiver was expecting an OW in response. The caregivers' continued prompting of OW forms in these cases highlights the mothers' engagement with these interactions, as she pushes the infant to go beyond simple repetitions of input speech to instead take part in a dialogue. As in Chapter 6, the routinized use of OWs in this way appears to be prompted by joint attention in book-reading activities, and this may be a specific feature of this particular context. In the toy-play situations, the OBJECT was often manipulated using movement (examples (3) and (4)), such that the toy became part of the interaction that was taking place; it seems that production of the OW alongside the manipulation of the OBJECT might have an advantage in early vocabulary learning. In a study observing the use of touch in infant-experimenter interactions, Seidl and colleagues (2015) found that the use of touch facilitated the learning of the corresponding words among infants as young as four months old. The authors state that "infants must be able to…track the co-occurrences of word form and referent" in running speech "for a word to earn a place in the early vocabulary" (2015, p.161). Their results highlight a role for infant-directed touches in particular, and we can propose from the results in the present study that the physical involvement of the OBJECT in the play situation – which was often manipulated by the infant while accompanied by the caregiver's OW production – may have a similar effect on OW learning.

Monaghan, Christiansen and Fitneva (2011) observed that arbitrariness was advantageous in language learning among adults, so long as context was available for those words. However, in a language acquisition setting, we have found that context is almost always present in the production of OWs, and so it seems unnecessary to consider infant language learning in the absence of context. Whether or not sound-meaning correspondences are found to be useful to infants in language learning tasks where context is not provided, in reality we have found that OWs serve to enhance caregiver-infant interactions within a specific contextual framework. However, we must question the reality of context here, as in this study we considered only those interactions which specifically involved OWs. If every utterance in the dataset were considered in terms of its context, it seems likely that we would find that the majority of caregiver-infant interactions take place within a relevant contextual setting: this is not a feature which is exclusive to OWs, but which can be found across all aspects of early word learning (Fernald & Morikawa, 1993; Woodward & Hoyne, 1999).

In this analysis, almost all OW production occurred during book-reading or toy-play activities – two situations which are reported in the language development literature to be beneficial to word learning. Rondal (1980) observed that lexical diversity was higher during book-reading activities than at mealtimes or during free-play situations, while Yont and colleagues (2003) found that book reading tasks elicited more focussed joint attention than toy-play tasks in interactions with 12-month-old infants. These authors also observed how infants were better able to direct their mothers' attention during toy-play tasks, as well as engaging more successfully in interactions relating to the activity in question. Furthermore, Soderstrom and Wittebolle (2013) report that both play and storytime situations provide the most varied linguistic input, although in their analysis both were very limited in the infants' daily routines, accounting for no more than 2% of waking activities. It seems that there is a consensus in the literature regarding the nature of these two activities in language development, highlighting how these situations may provide optimal learning conditions for young infants. The fact that OWs are generally produced within a setting that has been shown to benefit language learning further supports infants' extensive use of these forms in early production.

7.6 Study 3 – Interaction in OW production

In Chapter 6 we observed how mothers produced OWs in syntactic frames, as if engaging in 'dialogues' with pre-linguistic infants. Examples of this were also identified in Study 2, where we observed the caregivers' OW-elicitation in the later sessions, with some indication of how OW and CW production might interact in early development. This evidence suggests a role for OWs in the development of social-pragmatic understanding, as we observed how infants' 'incorrect' responses, although more adult-like with the production of the CW form, were not accepted as 'correct' by the mothers. This led to the establishment of turn-taking routine involving independent interactions with OW and CW forms.

In this study we will observe these routines in more detail, in order to understand how interactions change as linguistic ability improves, in line with Vygotsky's (1978) 'zone of proximal development'. This analysis aims to further clarify the role that OWs play in turn-taking routines, continuing from the observations in Study 2 which suggested that

OWs provide more than just lexical support to infants in the early stages of word production.

7.6.1 Analysis

OW and CW exchanges will again be considered, focussing this time on the specific interactions that surround production. This will reflect the use of OWs and CWs in infant-caregiver dialogues. Exchanges will be analysed in relation to five interaction categories, detailed in Table 7.4.

Table 7.4: Definition of interaction types used.

Interaction type	Definition	Example
SPECIFIC PROMPTING	The (attempted) elicitation of a specific OW or CW exchange.	*What does the cow say?* *What says moo?*
EXTERNAL PROMPTING	The use of questions to prompt word production in general, which leads to the production of an OW or CW exchange.	*Where's the doggie?* *What's the sheep doing?*
LABELLING	The production of OWs and CWs in relation to an image or toy without any obvious expectation of interaction.	*Look, doggie…woof woof*
RESPONSE	The production of an OW or CW exchange as a response to specific/external prompting or an external stimulus.	MOT: *What does the cow say?* **CHI: *Moo*** *Bump* (in response to infant falling)
REPETITION	Direct repetition of an OW or CW.	CHI: *Doggie* MOT: *Doggie, that's right*

All of these interactions constitute OW or CW exchanges between the infant and caregiver, which in most cases can be initiated by either of the two. In certain examples there is a clear role for each participant in the exchange, most notably in the case of SPECIFIC PROMPTING, where the caregiver is invariably attempting to prompt the infant towards the production of a particular word form, usually an OW. In total, 1,076 exchanges in the dataset were coded as one of these five interaction types – sixteen exchanges did not have a clear interaction type and were omitted from the analysis, while no interaction was considered for situations involving book reading or singing if the caregiver was producing a pre-specified text or song.

7.6.2 Results

Overview

LABELLING, RESPONSE and REPETITION are found throughout both the caregivers' and infants' data, while neither SPECIFIC nor EXTERNAL PROMPTING occur in the infants' production. The distribution of interaction types across infants and caregivers is shown in Figure 7.7.

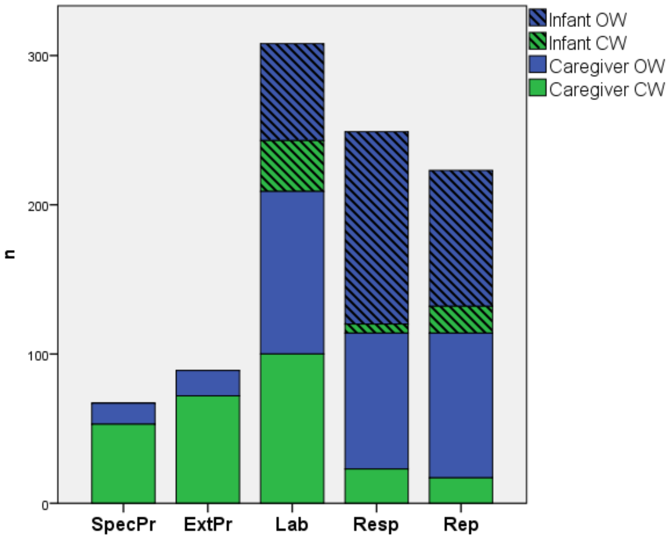

Figure 7.7: Interaction types across infant and caregiver production.

Here we can see that the use of SPECIFIC PROMPTING is the least common interaction type in the caregivers' outputs. This is mainly limited to the first two sessions: 81% of instances of SPECIFIC PROMPTING occur in Sessions 1 and 2, while LABELLING, RESPONSE and REPETITION (none of which involve explicit prompting of word production from the infant) dominate across Sessions 2 and 3. Production of the interaction types across sessions is shown in Figures 7.8a and 7.8b.

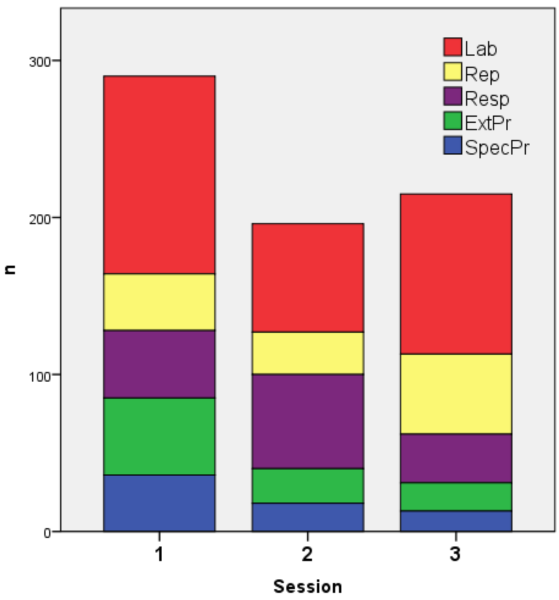

Figure 7.8a: Interaction types across caregivers.

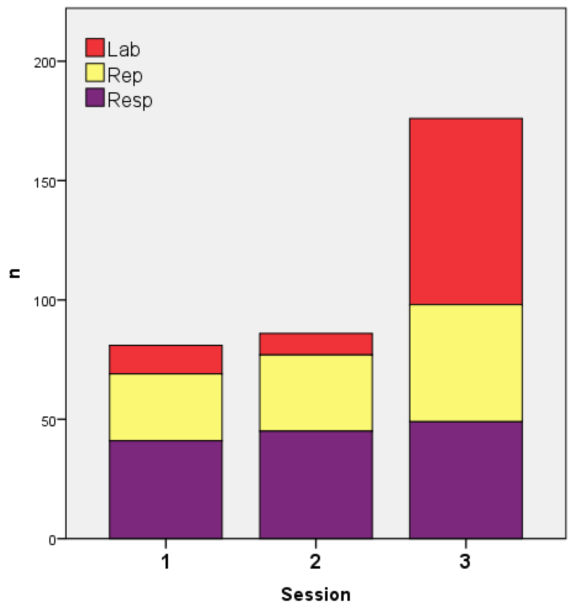

Figure 7.8b: Interaction types across infants.

The five individual interaction types will be analysed more closely in order to determine the specific nature of these exchanges between the caregivers and their infants. RESPONSES will not be discussed individually, but will be considered in relation to the use of SPECIFIC and EXTERNAL PROMPTING.

Specific prompting

All instances of SPECIFIC PROMPTING follow the highly conventionalised structure of *what does the [CW/OW] say?* with some modifications, including continuations of these phrases such as *what does a sea lion say?...how about a frog?* (Naima, Session 2). The conventionalised structure of this interaction leads to a large majority of CW exchanges here (82%), though we do find some examples of OW exchanges throughout the analysis, such as *what says moo?* (Alex, Session 3). The majority of these interactions are found in the American data: French caregivers account for only 15 (20%) of all the instances of SPECIFIC PROMPTING in the dataset, and all of the 36 instances in Session 1 are found in the American data.

A sharp decline can be observed in the use of SPECIFIC PROMPTING over time (Figure 7.8a), and this contrasts with the infants' RESPONSES to this interaction type: Figure 7.8b shows an increase in infants' use of RESPONSES over the course of the three sessions. A closer analysis of the data shows that in Session 1, infants respond to SPECIFIC

PROMPTING with the correct corresponding OW in 53% of instances, while by Session 3 the proportion of appropriate RESPONSES has risen to 85%: the caregivers' use of SPECIFIC PROMPTING decreases as the infants become more able to correctly respond to these cues. Advances in language ability do not lead to a change in use of SPECIFIC PROMPTING to prompt CW production instead of OW production; instead, this interaction type fades from caregiver speech. This was modelled in R using linear mixed-effects regression. Number of tokens was included as a dependent variable, with type (SPECIFIC PROMPTING vs. RESPONSE) and session as fixed effects, subject (dyad) as a random effect and by-subject random slopes for the effect of type. No effect was found ($\chi 2$ (1) = 2.96, p=.085), however, a Pearson product-moment correlation showed a significant relationship between caregivers' use of SPECIFIC PROMPTING and infants' RESPONSES across all three sessions (r=.486, n=24, p=.016). Furthermore, a more detailed analysis also showed a significant correlation between SPECIFIC PROMPTING in CW exchanges and infants' use of OW exchanges in RESPONSE (r=.41, n=24, p=.046). Here we see evidence to show how the caregivers' use of SPECIFIC PROMPTING directly affects the infants' overall OW production.

External prompting

EXTERNAL PROMPTING is also used most frequently in Session 1 (see Figure 7.8a), and decreases as infants' use of RESPONSES increases. Examples (11) to (13) below demonstrate the use of EXTERNAL PROMPTING (highlighted in bold) in interactions, where we can observe how caregivers follow EXTERNAL PROMPTING with SPECIFIC PROMPTING in order to continue a dialogue.

(11) Session 1: Lily

MOT: **is that a bear?**
MOT: what does a bear say?
CHI: arr !

(12) Session 2: William

MOT: **what's, what's that?**
CHI: dogs.
MOT: that's a dog, that's right, what's a dog say?
CHI: woof woof.

(13) Session 3: Nathan

MOT: **qu'est ce qu'il y a d'autre?**
what else is there?
CHI: meuh.

In Session 1, 71% of exchanges involving EXTERNAL PROMPTING are met with no RESPONSE from the infant, decreasing to one third of exchanges by Session 3. As found in the use of SPECIFIC PROMPTING above, by the final session the infants are responding with an appropriate OW or CW exchange in the majority of cases (67%), but by this point the use of EXTERNAL PROMPTING has diminished from caregiver-infant interactions (Figure 7.8a). As with the analysis of SPECIFIC PROMPTING, no significance was found when caregivers' use of EXTERNAL PROMPTING was modelled in relation to infants' RESPONSES with linear mixed-effects regression (p=.18), but here no correlation was found between these two factors either (p=.114). However, when external and SPECIFIC PROMPTING were considered together, a significant correlation was found between caregivers' overall use of prompts and infants' RESPONSES (r=.446, n= 24, p=.029). Again, we find evidence towards an explicit influence on infant OW production through the use of prompting in infant-caregiver interactions.

Labelling

LABELLING is the most common of the five interactions discussed here, accounting for 38% of all exchanges in the dataset. This interaction type does not have much impact on the infant output until Session 3, when it increases more than eight-fold from the previous session (Figure 7.8b). This appears to demonstrate independence on the part of the infant: LABELLING, in contrast with RESPONSE, reflects unsolicited word production, and an increase in this interaction-type over time indicates the infants' advancing linguistic abilities. Examples of infants' use of LABELLING in early CW exchanges are shown in examples (14) to (15).

(14) Session 1: Lily

MOT: **wolf.**
[SIT: looking at book]
CHI: **doggie.**
MOT: does the wolf look like a doggie?

(15) Session 2: William

```
MOT:   let's come sing what the wheels on the bus do.
[SIT:  playing with toy car]
MOT:   bvoom.
CHI:   cars.
```

Examples from Session 3 show that this approach continues over time: here the infants produce both OW and CW exchanges, with a greater variety of words which are often not quite appropriate for the given context.

(16) Session 3: Nathan

```
MOT:   ça c'est quoi?
       what's that?
[SIT:  playing with toy animals]
CHI:   woua woua.
       woof woof.
MOT:   un petit oiseau.
       a little bird.
CHI:   meuh.
       moo.
```

(17) Session 3: William

```
CHI:   swan.
[SIT:  looking at book]
MOT:   swan, right.
CHI:   duck.
MOT:   yeah, it kind of looks like a duck too, doesn't it.
CHI:   quack quack.
```

Here the infants' use of LABELLING appears to reflect a growing linguistic independence as they come to terms with their vocabulary, finding ways to associate their production capacity with the interaction taking place: note the obvious links between *wolf* and *doggie* in (14), *bus* and *cars* in (15), animals and animal sounds in (16) and *swan* and *duck* in (17).

Repetition

REPETITION is used most often in Session 3 across both infants' and caregivers' data, and for both groups this interaction type appears mainly in the production of OW exchanges (83%). Across the three sessions, very strong correlations can be found between the caregivers' use of REPETITION and that of their infants (Session 1: $r = .956$, $n= 8$, $p=.000$; Session 2: $r = .969$, $n= 8$, $p=.000$; Session 3: $r = .884$, $n= 8$, $p=.004$). We also see an increase in REPETITION across the three sessions: as the infants' vocabularies increase, so too does the use of REPETITION across the infants' and caregivers' data.

Patterns occur in the use of REPETITION in infant-caregiver interactions which can be found across dyads and over the span of this analysis. These patterns are demonstrated in examples (18) and (19), where we see the combination of SPECIFIC PROMPTING, EXTERNAL PROMPTING and LABELLING to elicit REPETITION by both the caregiver and the infant (exchanges involving REPETITION are highlighted in bold).

(18) Lily: Session 1

```
MOT:   what does a doggie say?
CHI:   woof.
        [ʌf]
MOT:   uff.
CHI:   woof
        [ʊf]
MOT:   woof.
CHI:   woof.
        [ʌf]
MOT:   woof
```

(19) Naima: Session 2

```
MOT:   you're a chicken.
MOT:   buck buck buck buck buck buck baga:h.
CHI:   buck buck.
        [ˈbæp ˈbæp]
```

A noteworthy change in the initiation of REPETITION exchanges can be identified over the course of the analysis. In Session 1, 84% of instances of REPETITION are initiated by the caregiver: that is to say that the vast majority of these exchanges involve the caregiver's immediate REPETITION of an OW or CW exchange produced by the infant (as shown in example (18)). In these cases, the caregiver appears to be consolidating the infant's early vocalisations with the accurate production of the adult target. By Session 2 only 53% of the REPETITION exchanges are caregiver-led, demonstrating the increase in infant-led REPETITIONS, whereby the infant identifies a familiar word produced by the caregiver and repeats it – seen in example (19): this appears to reflect a growing linguistic independence amongst the infants in this session. By Session 3, caregiver-led interactions increase to a slightly larger majority (59%), but of course, by this point, the infants are producing a much larger variety of word forms (see Figure 7.2); the increase in caregiver REPETITION here may be their attempt to support the infants' more varied word production.

7.6.3 Discussion

Here we have identified the changing role of OWs and CWs in turn-taking routines over the course of the three sessions. A number of trends can be found in the data that reflect a conventionalised approach to OW and CW production, as SPECIFIC or EXTERNAL PROMPTING is used to elicit OW production, followed by a cycle of REPETITION as the infant is encouraged to continue producing forms from their limited lexicon. Over time the caregivers elicit word production less often, as the infants becomes more capable of producing OWs independently. This gives way to caregivers' increased use of CW forms in Session 3.

These exchanges have been found to have their own independent roles in caregiver-infant interactions: SPECIFIC PROMPTING, for example, is used consistently with CWs (*what does the cow say?*) while REPETITION is much more common amongst OWs. Again we come back to the conventionalised use of OWs, whereby the routinization of caregiver-infant interactions specifies the contexts in which OW and CW exchanges are produced. This is also dependent on what the infant can (and does) say at any given point in time.

We have observed how both OWs and CWs are used in turn-taking routines, and how this may prompt infants' social-pragmatic understanding over time. In this respect, OWs have been found to have multiple roles, as they allow for the development of both lexical and prosodic features of turn-taking interactions. Examples (6) and (7) (Study 2) demonstrate the caregivers' reluctance to accept REPETITIONS as valid responses to their prompts, instead pushing the infants to provide the correct content in answer to their questions. Here we see how OW-CW pairings offer a rather exceptional opportunity in these contexts, as the semantic proximity of the forms as well as the extensive lexicon that they present (for example *what does an owl say?...what does a sea lion say?...what about an elephant?* – Naima, Session 2) can provide ample material for early 'proto conversations'. In turn, these allow for developments on a prosodic level, in line with Jaffe and colleagues' (2001) observations of infants' ability to partake in turn-taking routines, with adult-like mastering of conversational tempo and pause duration, amongst other features. Indeed, the authors observed that adult-infant dialogues were more coordinated than adult-adult dialogues, which is posited to be due to restrictions in the infants' language ability. Rochat (2001) comments on how this coordination of interactions enables infants to "gain the experience of rhythms of vocal and turn states as well as the coconstruction of novel sequencing or temporal parsing of auditory events" (p.140).

We have also seen further evidence for the ZPD here, continuing with observations from Study 1 where we remarked on lexical changes taking place over time in both the infants' and caregivers' OW and CW production. Here we have observed the changing nature of interactions, as caregivers move away from SPECIFIC and EXTERNAL PROMPTING just as the infants begin to respond appropriately to their prompts. Here we posit that this points to the caregivers' consistent calibration of their own production in line with that of their developing infants, and thus the shift away from prompting may push infants towards more independent word production over time. As infants gain an increasingly diverse vocabulary, they also become more equipped to take an independent role in interactions with the caregiver, moving away from the routinized structures that have been observed in earlier dialogues and towards more adult-like vocalisations: by Session 2 we see that REPETITION interactions are led by the infants just as frequently as the caregivers, showing a marked change from the predominantly caregiver-initiated interactions of Session 1.

Yont and colleagues (2003) propose that onomatopoeia merely provide "key words in game-like routines rather than…true discussions" (p.448), but here we have shown how these 'game-like routines' may facilitate the development of skills which lead to the ability to take part in more sophisticated dialogues. The conventionalised nature of OW forms has indeed been confirmed through this analysis, thus supporting observations from Chapter 6 where a social-pragmatic role for OWs was suggested. The routinization of OWs only points to their exclusive role in early interactions, whereby OWs and CWs can be drawn upon simultaneously to create turn-taking dialogues with adult-like sequencing, within the constraints of the developing lexicon.

7.7 Study 4 – Wildness in OW production

The final analysis in this chapter focuses on phonological wildness, as discussed in detail in Chapter 4. We observed the use of idiosyncratic sound effects in Annalena's OW production in Chapter 3, but here we will perform a more in-depth analysis of wildness in infants' use of OWs, observing how this corresponds to the caregivers' own OW production. Consequently, we will consider how the use of wildness shapes dialogues in early interactions, and whether this may provide any facilitative advantage to word production and interaction in early development.

While no effect was found for wildness in Chapter 4, infants showed longer fixations to target upon hearing stimuli with a higher pitch. Word-specific prosodic modifications were observed in the previous chapter which determined whether certain OWs were produced at a high or a low pitch, with a wide or a narrow pitch range: *toot toot* was produced with a monotonal high pitch to imitate a train, for example, while Figure 6.5c showed a long vocalic segment with descending pitch in the production of *neigh*, with a 'trembling' effect to imitate the sound in question. The combination of these findings suggests that infants pay more attention to those OWs with the more salient wild features. Similar characteristics are assumed to feature in this dataset, and so OWs will be analysed impressionistically for prosodic modifications, as well as more idiosyncratic wild 'effects'. This returns us to the iconicity debate from Chapter 4, where we called into question the precise nature of iconicity in OWs, which may be manifested in wild features rather than being inherent in all OW forms. An observation of the use of wild features (both prosodic modifications and the use of word-specific 'effects') in this analysis will allow us to further clarify the role that wild and tame OWs may play in early language development.

7.7.1 Analysis

Each of the exchanges in the dataset was coded for features of wildness: a word was judged to be produced with wildness if prosodic modifications or extra-phonetic features were used in its production, either to render the form more 'realistic' or to represent the sound or object in question. Characteristics of wildness include typical features of IDS – PITCH modifications (high or low pitch, or distinctive pitch modulations), VOWEL or CONSONANT lengthening and a noteworthy use of RHYTHM – as well as idiosyncratic 'special effect' features which were found to be typical of specific exchanges. For example, some instances of OWs *neigh* and *baa* are recorded as being produced with a 'trembling' voice (rapid opening and closing of the glottis to make a loud vibrato sound), *quack* as being 'nasalised', *woof* with a 'growling fricative' and *vroom* with a 'bilabial trill' (as in [ʙːːɹʊmː]). In all of these cases we account for wildness as being manifested in a specific EFFECT which is typical of the onomatopoeic word in question.

7.7.2 Results

Overview

Of the 1,159 exchanges (including 26 OW-CW combinations) recorded in the dataset, 287 were coded as containing wild features, accounting for 36% of all OW exchanges (n= 280). The majority of these (71%, n= 194) are produced by the caregiver. Figure 7.9 shows the proportion of all OW exchanges produced with wildness by infants and caregivers across the three sessions. Pearson product-moment tests reveal a strong correlation between the caregivers' use of wildness in OW production and that of their infants in Session 1 (r = .805, n= 8, p=.016): those caregivers who produced the highest number of wild OW exchanges had infants who also used the most wildness in their OW production here. The same test showed no correlation in use of wildness across dyads in either Session 2 (r = .468, n= 8, p=.242) or Session 3 (r = .568, n= 8, p=.142).

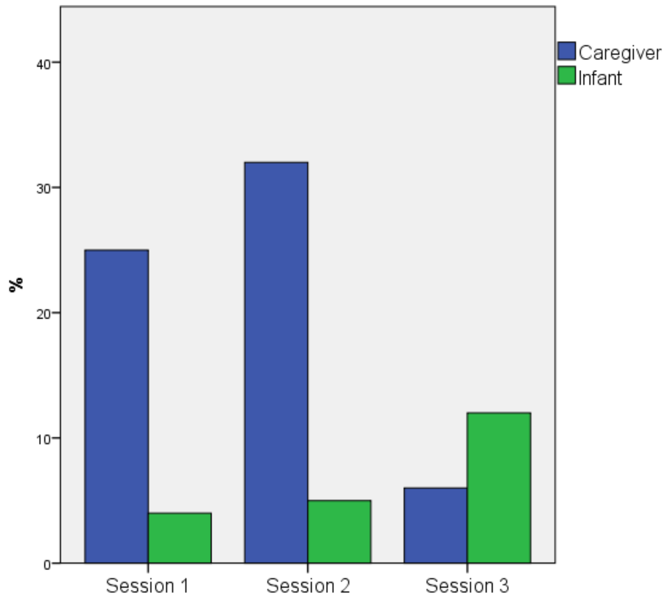

Figure 7.9: Use of wildness in production of OW exchanges.

Figure 7.9 highlights some inconsistencies in the use of wildness over time. In Sessions 1 and 2 we find a much higher use of wild features in the caregivers' OW production than in Session 3, when this decreases by more than half. In contrast, the infants' use of wildness in OW production more than doubles between Sessions 2 and 3, as they are found to be producing many more prosodic and extra-phonetic features alongside their OW exchanges. Nevertheless, it must be noted that wildness is produced in the minority of exchanges throughout the analysis, suggesting that this is consistently a peripheral feature of OW production.

In cases where CWs are produced with wildness, the same features tend to be used as would be expected in production of the equivalent OW – examples can be seen in (20) to (23) below, with wild forms highlighted in bold.

(20) Alex: Session 1

MOT: **I think I can I think I can I think I can I think I can I think I can I think I can. {RHYTHM}**
CHI: **chugga chugga chugga chugga**
[dɪgə dɪgə dɪgə dɪgi] {imitates MOT's RHYTHM}

(21) Lily: Session 1

MOT: chicken.
MOT: **boc boc boc boc boc.** {high PITCH}
MOT: **boc boc boc.** {high PITCH}
MOT: **chicken.** {high PITCH}

(22) Naima: Session 1

MOT: he's gonna ride on the ball.
MOT: **boing, boing, boing.** {RHYTHM + PITCH}
MOT: **there he is, on the ball.** {RHYTHM + PITCH continue from previous exchange}

(23) Théophile: Session 2

MOT: **miam miam miam miam.** {EFFECT: growling}
yum yum yum yum
[SIT: pretending to eat infant]
MOT: hum **j'ai faim moi.** {EFFECT: growling}
hmm I'm hungry

When combination exchanges are considered, trends in the use of wild features in OW and CW production remain. Fourteen of the 26 combination exchanges are produced with some wildness, but in all cases this occurs in the OW form but not in the equivalent CW. That is to say that wildness is not produced at an utterance level, but is instead dependent on the word in question. This is consistent across varying syntactical structures, as whether the CW occurs before the OW (as in *a sheep, baa* – Lily, Session 2) or vice versa (*woof woof, where's the dog?* – William, Session 1), we still find an OW-CW distinction in the use of wildness. This demonstrates that wild features are specific to the form in question, and so when CW exchanges are produced with wild features – though this only occurs on seven occasions in the entire dataset (see example (23)) – this is not merely an extension of wildness from a nearby OW, but instead can be assumed to be purposeful on the caregiver's part.

Features of wildness: caregivers

We see consistency in the type of features used across sessions. All caregivers use some wildness in the production of OWs over the course of the analysis, with seven of the eight producing wild OWs in both Sessions 1 and 2. By Session 3 the use of wildness has decreased on an individual as well as a group level: here five of the eight caregivers produce OWs with wild features. In all three sessions most OWs are produced with a single wild feature, which in the majority of cases (55%) is PITCH. This is important to note, since we observed a significant effect of pitch on infants' responses to OWs in Chapter 4. Figure 7.10 shows the characteristics of the prosodic features used in wild OW production over time, where we can see that pitch is by far the dominant feature across the course of the analysis, followed by VOWEL lengthening.

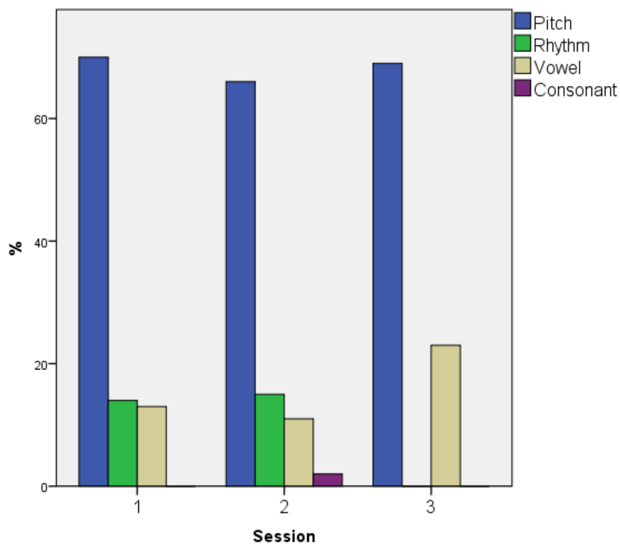

Figure 7.10: Caregivers' use of prosodic features in wild OWs.

Alongside the prosodic modifications that we would expect to see in IDS, we also see the extensive use of EFFECTS, which holds across the three sessions (Figure 7.11). Indeed, the proportion of OWs produced with specific special effect features remains almost constant even as the use of wildness in OW production decreases (cf. Figure 7.9). In Figure 7.11 we also see how the combination of EFFECTS with other salient prosodic features accounts for between 20-25% of all wild OWs in the first two sessions.

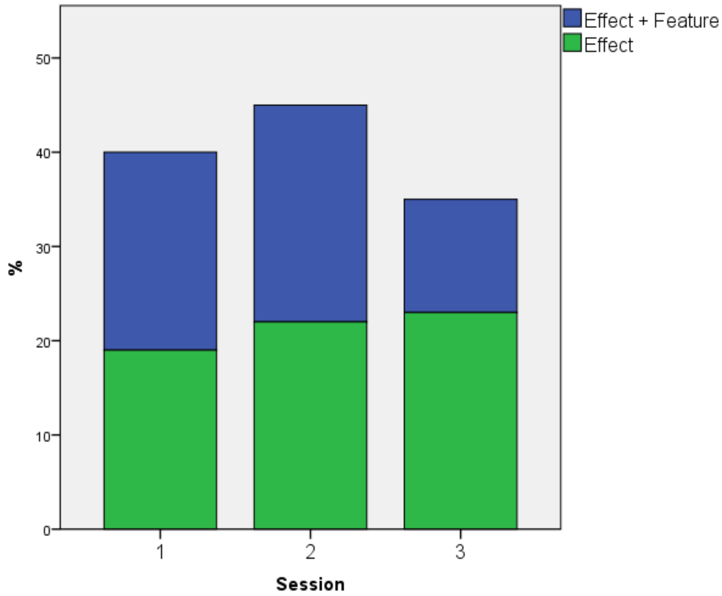

Figure 7.11: Caregivers' use of effects in wild OWs.

It seems that the use of wild features in the infants' input is highly idiosyncratic, and in up to 45% of cases these are specific to the OW being produced, occurring alongside the general features of IDS that we have already observed in Chapter 6. Furthermore, up to 23% of OWs are produced with an EFFECT in combination with one of the typical prosodic features of IDS detailed in Chapter 6 (Figure 7.11). OWs are thus often presented to infants with a range of prosodic modifications, both particular to IDS and specific to the form in question. However, as the majority of OWs are produced in a tame manner, we can assume that infants hear both wild and tame forms in the input.

Features of wildness: infants

Wildness in the infants' output is much more difficult to define, as in some cases it is unclear whether an infant's production of a word includes wildness, or if there are simply some articulatory limitations which bring about the production of features such as growling or consonant/vowel lengthening, for example. Eighty six (29%) of the infants' OW exchanges are produced with wildness, of which ten exchanges are recorded as being direct imitations of the caregivers' OW production: this indicates that infants are sensitive to caregivers' use of wildness, which leads them to use the same features in their own production of OWs. As shown in Figure 7.9, the production of wild features in the infants' outputs is much lower than that of the caregivers. However, trends can still be observed across the dyads. As with the caregivers, the most frequently used feature in the infants' wild OWs is PITCH, accounting for just under half (48%) of all wild tokens across the three sessions. Similarly, the second most common feature is VOWEL lengthening, constituting 20% of all wild OWs. Figure 7.12 shows both of these features to increase over time, as the number of wild OWs in the infants' inputs also increases. We also see that the infants make very little use of RHYTHM, and no use of CONSONANT lengthening.

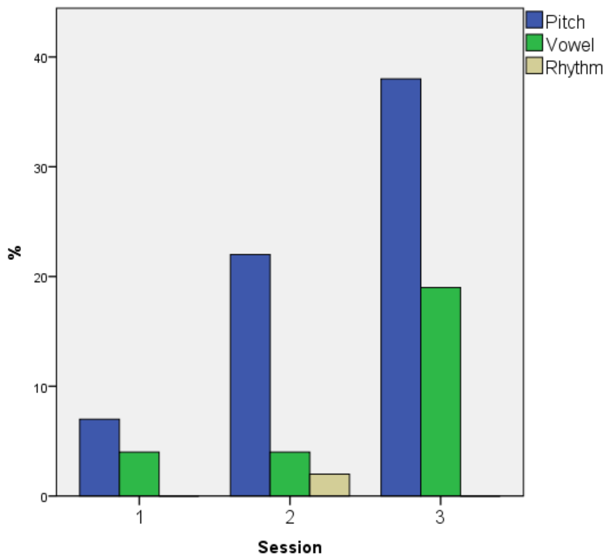

Figure 7.12: Infants' use of prosodic features in wild OWs.

When we observe the infants' use of EFFECTS we see that many idiosyncratic features are used in Session 3, but hardly at all in the first two sessions (Figure 7.13). Furthermore, the combination of EFFECT + feature is more common than the use of EFFECTS alone: 47% of the EFFECT + feature combinations in Session 3 also include PITCH, while a further 47% represent the only examples of CONSONANT lengthening in the infants' data (all produced by Théophile in imitation of a car). The final 6% constitutes a three-way combination of EFFECT + PITCH + VOWEL, again produced by Théophile.

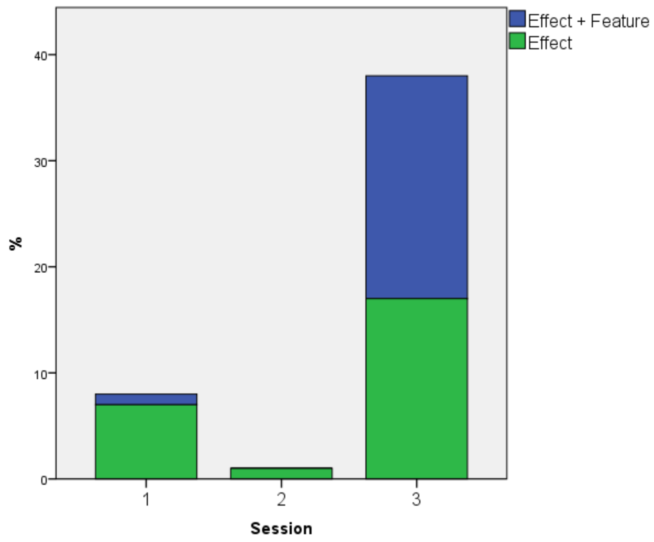

Figure 7.13: Infants' use of effects in the production of wild OWs.

Although wild OWs occur less frequently than tame OWs in the infants' production, it is clear that the use of wild features has a functional role in the infants' early output. Examples (24) to (26) below show how rudimentary the infants' early vocalisations are on a phonological level, but even so these are effective in communicating meaning owing to the accompanying wild features, which make them lexically distinct from one another.

(24) William: Session 1

MOT: What's this one?
CHI: **[m̥ m̥ m̥] {RHYTHM}**
MOT: Oh that's right (.) choo choo

(25) Naima: Session 2

MOT: What does an owl say?
CHI: **[ʔʌ ʔʌ ʔʌ] {EFFECT: growling + low PITCH}**
MOT: An owl not a dog
CHI: **[ʔʌ ʔʌ ʔʌ] {EFFECT: whisper + high PITCH}**

(26) Nathan: Session 3

MOT: Il fait miaou
He goes meow
CHI: **[a:] {PITCH: high PITCH + modulation}**
MOT: Oui
Yes

Example (25) provides a particularly striking demonstration of how wildness may be used to convey meaning in early interactions. Here Naima can be seen to produce two differing OWs – one identified by the mother as a dog, the other as an owl – which both have the same phonological structure in the infants' output. However, the changes in prosody across these forms, including PITCH changes and voicing EFFECTS, distinguish the two OWs as being separate, and lexically meaningful in their own right. This highlights how infants can use prosody in the absence of segmental stability in the production of OWs, thus allowing the infants to draw upon their linguistic resources when taking part in dialogues with the caregiver.

Individual differences

Across the dyads, wide discrepancies can be found in the individual use of wildness: some infants are exposed to many sound effects from their caregiver, while others hear a majority of tame forms. Similarly, there is no consistency across the dataset in infants' production of wildness with particular OWs; variability occurs both across and within sessions for each infant's use of wild features. For example, William produces numerous variants of *woof woof* in Session 1, two of which are produced with a distinctive PITCH modulation, two with a growling EFFECT and one with a combination of low PITCH, a growling EFFECT and a distinctive RHYTHM. In Session 2 Naima produces three differing forms of *bawk bawk* (to represent a chicken): five tokens are produced with no wildness, nine with a high PITCH, and one with a combination of high PITCH and a nasalisation EFFECT, which accurately imitates the mother's production of this OW. Naima also produces the OW *woof woof* with no wildness in Session 2 and with both a high PITCH (n= 2) and a growling EFFECT (n= 1) in Session 3, while Théophile alternates between a growling EFFECT and no wildness in his production of *vroum* 'vroom' in all three sessions. It seems that wildness is not phonologically specified in infants' production of individual OW forms, and instead can occur in free variation with the tame equivalent. In this case we must ask what prompts the infant to produce wild OWs, and whether this is specific to particular interactions or perhaps influenced by the caregiver's OW production.

Wildness in interactions

When turn-taking interactions between the infant and the caregiver are observed it is clear that the features used in the input have a strong effect on the infants' outputs. As observed in Study 3, repetition routines between the infants and caregivers are used extensively throughout the dataset, and notably in the production of OWs. We have also noted correlations between the infants' use of wildness and that of their caregivers in Session 1. Here, we consider only those interactions with repetition of OW exchanges between infant and caregiver (as shown in example (27), where exchanges with OW repetitions are highlighted in bold), in order to determine how wildness is used in these routines.

(27) Lily: Session 1

```
MOT  where's the doggie?
MOT  he's in here.
MOT  woof.
MOT  woof.
CHI  uff.
     [ʌf]
MOT  uff.
CHI  uff.
     [ʌf]
MOT  uff.
```

OW exchanges which were produced in isolation, or which were continuous repetitions by the infant not involving the caregiver (and vice versa) were not considered. This left 96 interactions for analysis, of which 30 were in Session 1 (n dyads = 5), 26 in Session 2 (n= 7) and 40 in Session 3 (n= 8), all of which involved at least two and as many as seven directly-repeated OW exchanges.

A striking consistency within W and T interactions is the first point to note here. All but five of the interactions (95%) involved exclusively W or T exchanges, and in many cases this appears to be specific to the particular dyads: Lily and Nathan both had predominantly T exchanges in their interactions (Lily: 82%, Nathan: 75%), while Alex, Marie, Théophile and William had a majority of W exchanges (Alex: 89%, Marie: 100%, Théophile: 75%, William: 65%). Anais and Naima had 52% and 43% T interactions, respectively.

We also see some direct imitations of prosody and wild effects in these interactions, where both caregivers and infants make use of imitation across all three sessions. Twenty three instances can be found of caregivers directly imitating the wildness in their infants' production across the dataset, with this occurring most frequently in Session 3 (n= 11; Session 1: n= 4, Session 2: n= 8). Four of the eight caregivers directly imitate the prosody of their infants' utterances (Alex, Lily, Naima and Théophile), with Alex's mother contributing the majority of these examples (70%). Perhaps more interesting, though distinctly less frequent in the dataset, is infants' direct imitation of their caregivers' use of prosody in wild exchanges. Only seven examples occur across the three sessions, but these are distributed across six of the eight infants (all but Anais and William), demonstrating the influence of the input on the infants' production of OW exchanges, and the infants' awareness of wildness as a feature in OW production.

7.7.3 Discussion

An exploration of the use of wild features in infant-caregiver interactions has confirmed Rhodes' (1994) claims that a wild-tame paradigm exists in the production of OWs, and moreover, that this may serve as a functional aspect of early infant communication. The production of wildness was identified as being both independent of and in tandem with the prosodic features reported to occur in IDS: in some respects, wild forms fit within the typical features of IDS, while we also observed the production of idiosyncratic extra-phonetic features which were specific to the OW in question. Rhodes does not specify whether his paradigm applies to child or adult language, but here we can confirm that, at least in interactions with infants, both sides of the wild-tame dichotomy exist in the outputs of both adults and infants. Moreover, a relationship was identified between the presence or lack of wildness in both the early input and the output in Session 1, as the correlations observed between infants' and caregivers' general OW production in Study 1 were found to be consistent with the use of wild features. This did not extend to the later sessions, however.

To a large extent the wild features applied in the dataset correspond to the features that are commonly reported in IDS. Pitch changes – incorporating high pitch, low pitch and modulations – were found to be dominant in the production of OW exchanges, as well as vowel and consonant lengthening, and use of rhythm in word production. None of these features are unusual in caregiver-infant interactions, especially in the early stages of word production, and so we witness nothing especially noteworthy or specific to OWs here (though, of course, OWs have already been shown to be prosodically more salient than CWs, which is a particularly noteworthy feature in itself). However, we do find many examples of idiosyncratic prosodic and extra-phonetic effects relating to specific OWs: growling sounds to refer to dogs or engines, for example, or 'trembling' intonations in reference to horses or sheep, as well as creaky voice in the production of *ribbit*, and snorting to imitate a pig – none of these features are reported as being part of IDS, and are particular to OW production.

The distinction between wildness and salience should be noted here, as it can be inferred that the use of wild features will render an OW more salient than its tame equivalent. In particular, we can assume that the more prominent use of pitch features in these forms is particularly important, based on findings from Chapter 4. However, while we can draw a legitimate hypothesis based on results from previous chapters, without an acoustic analysis of the data in these recordings we are unable to provide a definitive judgement on this aspect of the data. The analyses taken in the present study are impressionistic, and therefore we are unable to objectively compare the salience of specific words. Nevertheless, evidence from these recordings allows us to posit that wildness may render OWs more salient in both the input and the output: the free variation of wildness and tameness in each infant's OWs has shown infants to be aware of this paradigm. Moreover, the infants were often found to draw upon it independently, through the use of wild features to distinguish phonologically-similar forms.

Perhaps the most revealing finding here is the use of wild features as meaningful elements of production. The addition of specific prosodic or idiosyncratic effects was shown to differentiate between one vocalisation and another – forms which would otherwise be indiscriminable if considered from an exclusively phonological perspective. Indeed, wild features enabled the transfer of meaning in infant-caregiver communications in the absence of phonological capacity, as the infants were found to produce simple vocal gestures – often a single vowel or consonant, or a CV syllable – with the use of prosodic and extra-linguistic effects to convey additional lexical information to the caregiver. This was first observed in Annalena's 'innovative' OWs in Chapter 3, and we can relate to the same functional use of special effect features in her data to overcome difficulties presented by limitations in her phonological capacity. With regard to the wild features used in this dataset, we observe how these adhere to the limitations of early production; use of pitch and vowel lengthening were by far the most common prosodic properties identified in infants' wild forms, which is unsurprising as these features are rehearsed from the very earliest vocalisations. Indeed, the manipulation of pitch is reported in the pre-linguistic stages of vocal play (Stark, 1980), while open vowel sounds are present from birth (Stark, 1978, 1980, 1989); in contrast, we cannot expect to observe consonantal stability until the phonological inventory is better established. However, we do see some segmental control in the production of effects, where consonant lengthening, as well as differing phonation types such as growling, was observed for specific segments. This was most notable in non-native phonemes such as the bilabial trill [ʙ] as well as sonorant fricative sounds such as [z̥] or [x], for example, which are produced as sound effects rather than phonological segments of the ambient language. Stark (1980) discusses the production of consonantal noises in pre-linguistic infants' outputs, and notes that "stops, clicks, friction noises and trills" occur in very early vocal production as "primitive precursors of the nonsonorants found in adult spoken languages" (p.78). Again, this returns us to Annalena's early production in Chapter 3, where she was found to produce certain segments, including non-native consonants, exclusively in OWs.

One point regarding infants' use of wildness which appears to contradict its advantage in early communication is the fact that wild features were used very rarely in Session 1, increasing over time to represent 38% of all OWs in Session 3. The proposition that wildness may be used functionally in the absence of phonological capacity would lead to the expectation that wild features would be used more consistently in the earlier sessions, when the infants' language capacities are yet to be established. This is not the case in this data, but the increase in wild features can be seen to parallel that of the infants' increasing OW production, suggesting that the infants continue to use wildness functionally over the course of the developmental period considered in this analysis. We have also observed an increase in the number of CWs by Session 3, pointing to a broader understanding of the use of OWs in interactions: as infants begin to consider OWs as sound effects, the use of wild features may become more prolific, as *meow* is used to designate the noise that a cat makes, as opposed to a sound produced in broad reference to cats.

The optional presence of wildness in both the input and the output is also important to note. All infants were exposed to wild and tame OWs in the dataset, and all were found to produce OW exchanges both with and without wildness. Although certain dyads appeared to prioritise either wild or tame OWs in their interactions, it seems that infants' representations of OWs include the free variation of wildness and tameness; seven of the eight dyads used both wild and tame OWs in their interactions. The use of both wild and tame forms demonstrates the generalizability of OWs in early perception and production, and perhaps also an understanding on the infant's part of the idiosyncrasy of these forms. However, wildness, or lack of wildness, was found to be consistent throughout individual caregiver-infant dialogues; in some cases this included up to seven separate exchanges, but even the longest repetition dialogues in the dataset showed consistency of W or T forms throughout. The variability between interactions combined with the consistency within interactions demonstrates the dynamic nature of OW production in terms of the use of wild features in early communication.

This analysis has confirmed that wildness (in free variation with tameness) is a valid feature of OW production, and can be found in both caregiver and infant speech. However, wildness was present in just over a third of all OW exchanges in the dataset, and so it seems that tame OWs provide the 'default' forms here. Once again we have observed a role for the input in this analysis, as the infants' use of wildness was found to be dependent on that of the caregiver, especially in the earliest session. Furthermore, the use of wildness to distinguish phonologically similar forms was particularly revealing, as this provides evidence for the claims made in Chapter 4, where we suggested that wildness may provide perceptual information over and above that of phonology – this is a feature that is exclusive to OWs, and has been shown here to be applicable to the early output, as well as to the input. Overall, it seems that infants draw upon wildness in OW production, taking advantage of the status of these forms to further early word production through the use of prosody and supra-segmental features in the absence of articulatory control.

7.8 General Discussion

All caregivers in this analysis produced OWs in all three recordings, and, with the exception of Anais, OWs were also recorded in all of the infants' outputs across the three sessions. Indeed, even in the first session, when some of the infants produced as few as six words, still OWs were found to play a role in the infant-caregiver dialogues observed here. This confirms the reality of OWs in early interactions, accounting for the consistent use of these forms both over time and across languages. Further to this, the four separate studies conducted in this analysis have each contributed to a more in-depth perspective of infants' acquisition of onomatopoeia in early language development. Individually, each analysis has provided further insight towards the main research questions in this thesis: Study 1 demonstrated a strong relationship between the caregivers' use of OWs and that of their infants, while Study 2 showed OW production to be motivated by stimuli in the surrounding environment. Study 3 outlined the various interactions that lead to OW production and how these take place between infants and caregivers over time, as language develops to allow the infant more vocal independence. Finally, Study 4 confirmed a role for wildness in OW production, which was found to be variable in both the infants' and the caregivers' use of OWs, while also being functional in the early output.

The interactions observed in this analysis can be assumed to provide the infants with an important opportunity for language learning, as in all cases these involved extended periods of focussed joint attention, with large amounts of linguistic input as well as contextual prompts from external stimuli. Indeed, alongside studies showing that interactions with the caregiver are indicative of an infant's language learning (Tomasello, 2003), the context of early interactions has also been reported as highly relevant to infants' language development, which can be related to the interactions observed in this analysis. Soderstrom and Wittebolle (2013) found that caregivers provide the most variable lexical input during storytelling and structured playtime activities, as infants were presented with a more lexically-diverse input, as well as the opportunity for focussed interaction with the caregiver. It is proposed that these contexts provide particularly salient learning opportunities, although they were found to constitute a very small proportion (1-2%) of the infant's overall daytime activities. This corresponds to the large proportion of OWs produced in the context of toy-play and book-reading activities in this analysis, as we see evidence of optimal language learning situations here. Furthermore, Rochat (2001) discusses the importance of toys and other objects in early interactions from a different perspective, describing how they enable infants to adopt conversational roles and create dialogues in the home. Evidence for this was found in Study 2, where both infants and caregivers were found to bring these external objects into their interactions, thus engaging more thoroughly with the words that were bring produced, while also creating imagined scenes and dialogues with the animals and objects in question.

Here we see more evidence for a role for onomatopoeia in early language learning, in relation to the acquisition of skills necessary to interact effectively with other speakers. Interactions such as SPECIFIC PROMPTING may be useful in establishing the rudiments of turn-taking in conversation from the very early stages of lexical development. These interactions provide infants with the opportunity to develop an understanding of turn-taking in terms of timing and rhythm (Jaffe et al., 2001), enabling participation in dialogues which are guided by the caregiver. OWs provide a functional sub-lexicon in early word production which makes such interactions possible, owing to the simple structures of these words which allow for a more flexible form in the output. Study 4 showed how wildness can be implemented here, since infants demonstrated an ability to use particular prosodic features to make words meaningful when they might otherwise be unidentifiable. It is not surprising, in this case, that OWs are drawn upon in early conversation in order to maintain interactions between infant and caregiver, as they function as expressive early protowords which are meaningful to both interlocutors, enabling functional 'protoconversations' (Rochat, 2001) which lead to further language development. Extensive use of repetition should also be discussed here, as this allows the development of articulatory precision in word production embedded in a turn-taking routine to provide further social pragmatic experience. The consistencies observed in the use of wild features further establish this point, as multiple repetitions in an interaction (sometimes up to seven instances of the same word in one dialogue) do not bring about modifications in the forms that would suggest creativity in the dialogues. It seems that these repetition routines serve to establish the infant's representation of a word form, in terms of both the perception and production of that form. Furthermore, Zimmerman and colleagues (2009) show that more extensive turn-taking routines as observed here are strongly associated with larger vocabularies in later development, suggesting that the to-and-fro of repetitive routines may be beneficial in the long term.

Another striking finding from these analyses was identified in Study 3, where turn-taking dialogues were found to follow specific patterns which were modified over the course of the analysis according to the infants' developing language abilities. Changes were observed over time which appeared to be determined by both the caregiver and the infant: as the infants' language abilities improved, their independence in production increased, and well-rehearsed production routines gave way to less structured and more variable interactions. A routinized structure to OW and CW production was initially observed in the previous chapter, when caregivers were found to present pre-linguistic infants with 'call and response' dialogues even though the infant could not take part, with the mothers fulfilling both sides of the 'interaction'. A continuation of this was observed in this analysis, whereby the same routines were used across the board in the earlier sessions, when the infants were linguistically less capable. However, once infants were able to take part verbally, the caregivers were found to move towards more sophisticated forms of interaction, again pushing the boundaries of the infants' language and communication abilities. The patterns here relate to Vygotsky's (1978) 'zone of proximal development', as caregivers were found to calibrate their speech to the infants' changing language abilities. Examples from the data show how this may promote language learning in these contexts, as the caregivers continually adapted their speech "to be *just challenging enough*" (Zimmerman et al., 2009, p.347, italics added), such that infants were able to take part in dialogues mediated by the mothers' prompts. Furthermore, Zimmerman and colleagues propose that, for the ZPD to function effectively, the adult must be in touch with their infant's developing language abilities, which can be maintained through continuing interactions with the infant. Here we are presented with another perspective on this data, whereby one could suggest that the caregiver-infant interactions observed in this analysis are not only productive in terms of the infants' language abilities, but also in terms of the caregivers' role in their infants' language development. Engagement through the routinized use of OWs and CWs may provide the opportunity for the caregivers to continually reassess their infants' linguistic capacity in relation to the input that they are providing.

Dialogues such as these are considered widely in the literature. Garrod and Pickering (2004) point out that language rarely takes place in a monologic form, and indeed that "it is through dialogue that humans learn to speak" (p.11); the processing load is shared between the two interlocutors in a dialogue situation (Garrod & Pickering, 2004), making these interactions easier for infants to engage in. However, infants in the early stages of language learning must come to understand the nature of language in conversation, as well as to acquire the lexical, grammatical and syntactic skills necessary for speech, and so dialogues are also essential for developing general communicative capabilities. In this analysis we have identified how OWs may facilitate this process, as routinization – observed in this study as well as in the previous chapter – is considered by Garrod and Pickering (2004) to be a necessary tool for simplifying production processes in all forms of conversation, and is found here to be important for infant-adult dialogues as well as the adult-adult dialogues to which the authors originally refer. Furthermore, Jaffe and colleagues (2001) found that infants as young as four months had already developed the required linguistic skills for turn-taking interactions, to the extent that they were able to engage with strangers in a laboratory setting as well as with their mothers in the home. The authors state that infants' engagement in interactions with the caregiver provides essential social pragmatic experience, which can then be extended in unfamiliar situations from a very young age. Indeed, this is the case even before infants are able to speak, and is particularly appropriate to OW production as the infant can make use of extra-linguistic and prosodic effects in order to communicate onomatopoeic protowords in the absence of an established vocabulary. Rochat (2001) refers to this as 'protoconversation', as infants gain experience of parsing speech and representing external events even before their linguistic capacity is sufficiently sophisticated to articulate these events through words. Indeed, in this analysis infants were found to participate actively in dialogues with the caregiver, not only as responders to their caregivers' prompts, but also through leading interactions and shaping both the content and even the prosody of the dialogue.

Throughout this thesis we have questioned the presence of OWs in infant-caregiver interactions: is infants' production of these forms motivated by their internal phonological features, believed to be advantageous in the early stages of language learning (Imai & Kita, 2014; Werner & Kaplan, 1963)? Or is infants' OW production driven by the caregiver, either through prompting on the adult's part or through the infants' repetition? Overall, the results from these four analyses show that the presence of OWs in early infant speech is driven by a number of factors, supporting findings from earlier chapters which have provided different perspectives on the multifaceted nature of these forms in early language development. Findings from this chapter go against Werner and Kaplan's (1963) theoretical perspective on infants' construction of the world through the imitation of sounds; in the majority of circumstances, OWs were produced in a context that did not involve the infants' direct vocal imitation of a natural sound (although this approach to OW production was identified in two infants' data). Furthermore, throughout the analysis we found a strong connection between the caregivers' use of OWs and that of their infant. This was found not only in the extent to which caregivers used OWs, but also in the context in which they used them and the interactions that were observed; infants whose caregivers produced OWs only in specific contexts also limited their use of OWs to these contexts, and this was consistent across stimuli such as books and toys. Théophile's data consolidates these findings, as his approach to OW production could be interpreted as supporting Werner and Kaplan's (1963) position regarding the creation of onomatopoeic protowords through imitation. However, his parents were found to have the same approach to OW production, pointing to the input as the main determiner – both qualitative and quantitative – of Théophile's OW use. We also observed a potential role for interaction types across the infant-caregiver dyads, as the extent to which the caregiver used SPECIFIC and EXTERNAL PROMPTING appeared to be reflected in the infants' OW production, at least in the first two sessions. These two cases present different aspects of caregiver influence on the infants' OW production: on the one hand it appears to be driven implicitly through experience of OW production in the input, while on the other hand we see how an infant's OW use may be explicitly motivated by the caregiver. Together this evidence provides a new perspective on infants' production of OWs, which does not correspond with the theoretical positions positing the implicit learnability of OWs (Imai & Kita, 2014; Werner & Kaplan, 1963). Instead, this leads us to propose that OW acquisition is determined externally, brought about by the caregiver's own OW production in interactions with their infant.

The establishment of sound-meaning links is one theme which comes up regularly in the discussion of onomatopoeia in language learning, and is central to our investigation in this thesis. Evidence from this analysis shows that onomatopoeia may indeed facilitate infants' understanding of forms and meanings in early development, but that this is brought about by the consistent pairing of onomatopoeia with relevant referents such as toys, actions or pictures in books, rather than any implicit iconic advantage in these forms. Indeed, OWs were paired with visual referents such as toys or pictures in almost all cases in this analysis, even more so than they were paired with their equivalent CW forms. It is therefore unsurprising that OWs are acquired early in development, owing to the presentation of such unambiguous form-meaning mappings in the early input. Furthermore, in many cases this is combined with the equivalent CW, resulting in a three-dimensional representation of a form whereby the object, its referent and a corresponding sound effect are presented together. This is in contrast with the 'mutual exclusivity hypothesis' (Markman, 1994; Markman et al., 2003), which has shown that infants assume a 'one object, one label' status in word learning. Indeed, the mutual exclusivity perspective would assume that this three-dimensional representation would be confusing for infants, but instead it opens up opportunities for establishing interactions between caregivers and infants with only the rudiments of a lexicon. The 'one object, two labels' status of onomatopoeia is shown here to facilitate word production, as well as establishing early conversations through well-rehearsed production routines.

It seems that the infants observed in this study are experiencing OWs in optimal conditions for language learning. Indeed, it is no surprise that we see such a prominent use of OWs in the early output if these forms constitute such an important feature of shared interactions with the caregiver. Not only do OWs provide opportunity for turn-taking routines, but they are also widely represented in books and through toys, presenting infants with a three-dimensional representation of these forms. It is well-established in the literature that joint interactions with the caregiver promote vocabulary development, and also that play and book reading activities assume an important role in supporting infant language learning. Pairing this with the routinized approach to OW production in infant-caregiver dialogues, which sits within the framework of a continually changing ZPD, we see here how OWs provide a unique opportunity in language learning: use of OWs incorporates multiple facets of the learning experience within interactions which appear to take place naturally between infants and their caregivers.

7.9 Conclusion

The separate studies in this analysis have allowed us to address some of the questions that remain from the previous chapters in this thesis, consolidating the evidence with real-world interactions which show the use of OWs in early language development. A strong link between the early input and the developing output has been identified, which allows us to consider OWs within the wider picture of language development: while constituting only a small subset of the lexicon, still OWs are a learned feature like any other, and their learning is dependent on the extent to which infants are presented with these forms, and the nature in which they experience them. Furthermore, OW-based interactions taking place between the infant and caregiver have been shown to contain multiple facets of early language experience which are established in the literature as essential to language learning. OWs thus appear to provide opportunity for learning in a lexical, syntactical and social-pragmatic domain.

8. Discussion

8.1 Overview

This thesis has attempted to address the absence of empirical research into infants' early production of onomatopoeia. In order to fill this gap in the literature, we have investigated the role of onomatopoeia in phonological development from three perspectives: production, perception and interaction. This "cumulative science" approach (Cristia, 2015) has brought us to a number of firm conclusions, which allow us to present evidence in contrast with the theoretical positions positing that the inherent iconicity in these forms is what leads to their abundance in early lexical development. In the introduction to this thesis we questioned whether onomatopoeia can viably be considered as iconic, and whether they should be treated alongside sound symbolism and mimetics in the literature. We suggested that onomatopoeia may in fact be no more iconic than any other words, and as such they should be acquired in the same way – namely, through the interaction of perception and production in early language experience. While to a certain extent this has been shown to be the case in this thesis, the six studies presented here have provided us with an alternative perspective, which allows us to consider onomatopoeia as both separate from and integrated with phonological development in perception, production and early interaction.

The main body of literature presents us with two perspectives on onomatopoeia: first, the more traditional view, most famously put forward by Werner and Kaplan (1963), which advocates the theoretical position of onomatopoeia as 'stepping stones' to a more adult-like lexicon. Werner and Kaplan (1963) assume onomatopoeia to be created by the infant as "vocal depictive forms" (p.101) as language unfolds ontogenetically, from the earliest reflexive vocalisations into full sentences. However, even a brief consideration of the early onomatopoeia produced by infants quickly contradicts this theory, since the majority of these forms are conventionalised, as in *woof woof*, and thus clearly learned from the input. The second perspective is explored in very recent research, which has largely considered onomatopoeia as part of the broader debate on iconicity. While there is of course good reason to class onomatopoeia as iconic, this also places these forms under the same rubric as sound symbolism, and works on the assumption that all iconic forms are fundamentally the same in infant language acquisition. Thus, the recent research showing that infants develop sensitivity to sound-form correspondences at a very young age (Asano et al., 2015; Ozturk et al., 2013; Peña et al., 2011) has led to the assumption that this sensitivity is also relevant to onomatopoeia, and that this must be advantageous to infants acquiring their earliest words (the 'sound symbolism bootstrapping hypothesis', Imai & Kita, 2014). While founded on different theoretical frameworks, these two approaches have two important factors in common: first, both assume that onomatopoeia work as a scaffolding mechanism into further language development; and second, neither has provided any empirical evidence with specific regard to onomatopoeia to support this. Nevertheless, this has provided the basis for a blanket assumption in the literature that onomatopoeia have an iconic advantage in language learning, which makes them easier for infants to acquire.

Here we present a new perspective on infants' acquisition of onomatopoeia, which provides insight into why we find so many of these forms in early infant lexica across so many languages. A range of different methodologies have allowed us to explore this question in detail, with perspectives from perception, production and early interaction combining to present a picture which is in fact unsurprising, and which fits well within our general understanding of how infants learn language. In many ways onomatopoeia have been shown to fit within the developing linguistic system just like any other word form, and so we can begin by concluding that these forms should not be overlooked in studies of phonological development. This is consistent across the three main aspects of this analysis:

- In *perception*, OWs were found to be significantly more salient than their equivalent CWs in terms of the prosodic features reported to facilitate speech segmentation and word learning in IDS (see reviews by Cristia, 2013 and Soderstrom, 2007). Infants attended longer to the forms produced with a higher pitch, suggesting that this may be a particularly important feature of OWs. Moreover, the common presence of reduplication in OWs (as well as frequent repetition) meant that infants were presented with more than twice as many

OW tokens than CW tokens, giving these forms a distinct advantage in terms of frequency (cf. Ambridge et al., 2015).

- In *production*, the simple shapes of OWs were found to be typical of the prosodic structures that infants tend to produce in early speech. Reduplication, consonant harmony and single CV syllables are all characteristic of OWs, and are also considered to be 'universals' of infant speech in the literature (Kent, 1992; Smith, 2004). Consequently, infants 'select' OWs (Vihman, 2015) in line with their phonological capacity in early development, as these forms often provide an ideal match to infants' preferred output patterns. On the other hand, the phonological structure of OWs tends not to be strictly specified, allowing for their flexible production without compromising on meaning. The use of prosody and wild features is also advantageous here, as it allows infants to communicate effectively in the absence of a stable phonological inventory.

- In *interactions*, OWs provide ample material for the prompting of early word production, within the framework of turn-taking routines. Routinized syntactic structures were found across the data, whereby caregivers first assumed the position of both interlocutors by producing both OW and CW (when the infants were unable to respond), followed by the prompting of OWs through rehearsed production routines (as in *what does the cow say?*). Eventually this led to more independent vocalisations on the infants' part, involving CWs as well as OWs. Caregivers were found to be pushing their infants' language development forwards at all stages, within what could be interpreted as a zone of proximal development (Vygotsky, 1978). In this respect, we observed a continuous move towards the more adult-like CW form, which was consistently produced in combination with the equivalent OW.

Together, these findings lead us to conclude that onomatopoeia have an advantage over many other word forms, as they are more salient in the early input and well-suited to the articulatory constraints of the early output. In this respect, it is unsurprising that OWs are found so often in infants' early words. However, the direct link between the caregiver's use of OWs in early interactions and the infant's eventual acquisition of these forms cannot be overlooked: as we would expect in all other aspects of phonological development, infants must experience OWs in the input if they are to acquire them. Thus, while we propose that there is an advantage for OWs in early language development, this is driven first and foremost by infants' experience of these forms. Furthermore, this advantage is not dependent on the inherent iconicity that these forms are assumed to possess, but is instead manifested in the perceptual, productive and interactional idiosyncrasies discussed above, which set them apart from the rest of the developing lexicon.

8.2 Perception-production interface

In Chapter 2 we observed how OWs fit within infants' preferred output patterns, which are selected from the motor routines produced in their canonical babbling, leading to the establishment of vocal motor schemes (VMS). VMS are defined by McCune and Vihman (2001) as "generalized action patterns that yield consistent phonetic forms" (p.673); the authors observed how infants with a number of VMS in their babbling went on to produce their early words in line with this well-rehearsed repertoire of consonants. Here we can see how the simple, often reduplicated, prosodic shapes of OWs may derive from canonical babbling, and how they may come to be established in infants' early words as meaningful lexical items.

Moreover, when we note the frequency of OWs in IDS, as well as their salience in the early input (Chapter 6), we can posit that infants' early production of OWs may be consolidated by their experience of these forms. This relates to the 'articulatory filter' (Vihman, 1993), or 'auditory-articulatory loop' (Stoel-Gammon, 2011), which "renders similar patterns [to the infant's own output forms] in adult speech unusually salient or memorable" (Vihman, 1993, p.74). In the early stages of production, infants with just one stable VMS have been found to show a preference for the consonants that they are able to produce (Majorano et al., 2014). However, once infants' production abilities become more advanced, they begin to attend more to those consonants that are less familiar (DePaolis et al, 2013; Majorano et al, 2014). In the case of OWs, the 'match' between the caregiver's production of these forms and the infant's own babbling patterns, combined with their salience in the speech stream, may give OWs a particular learning advantage in early language development.

The articulatory filter may also be relevant to the use of wild features in OW production. Infants have been shown to master prosodic properties of their language such as pitch and rhythm (Levitt, 1993) earlier than the segmental properties, and evidence for a language-specific prosodic bias in the earliest stages of both perception and production is well-established (Boysson-Bardies et al., 1984; Mehler et al., 1988). Some degree of prosodic control is available to infants prior to any segmental capacities, and it is even proposed that infants begin to master vocalic speech-like elements through crying, including respiratory timing and prosodic features such as pitch, intensity and rhythm, as well as the use of articulatory gestures (Stark, 1989). Such features of 'vocal play' (Stark, 1980) constitute infants' earliest experience of production, and so caregivers' extensive use of salient prosodic features in the production of wild OWs may also draw infants' attention to the prosodic aspects of their own output. Consequently, this may bring about the infants' imitation of these wild features through their own mastery of prosodic modulation. Indeed, from the age of only three months, infants are able to match pitch features in the input in terms of both fundamental frequency (Kessen et al., 1979) and pitch modulations (Papoušek & Papoušek, 1989). With this in mind, we posit that infants' very early ability to imitate features of pitch, as well as vowel quality (Kuhl & Meltzoff, 1982), facilitates their early production of OWs. Attention to the distinct prosodic effects used in mothers' production of OWs, combined with the infants' own ability to match prosodic features in the input, may enable infants to attempt these forms even before the earliest phonological segments have been stabilised. This could also be related to infants' tendency to imitate sounds in the surrounding environment, as observed in Annalena's data in particular (Chapter 3); infants' attention may be drawn to the features of pitch in both vocal and environmental sounds as a match to their own vocal capacities in the pre-linguistic phase.

From this analysis, supported by evidence presented in this thesis, we can return to Jakobson's (1941/80) claims regarding the stabilisation of an early phonological system through the production of OWs. Indeed, their unprescribed phonological structure allows infants to produce these forms even with the most rudimentary of vocal capacities. Infants are able to communicate meaning through the imitation of wild features: the production of rudimentary vocal gestures combined with pitch control and modulation. In Chapters 3 and 7 we observed infants' production of wild forms, whereby idiosyncratic effects and non-native phonemes were produced in onomatopoeic protowords. The fact that these simple forms could be used as meaningful lexical items in conversations with the caregiver demonstrates how the phonological flexibility of OWs may be especially important in early development:

again, these forms present infants with an opportunity to practice word production within the limited confines of an early phonological system. As Jakobson (1941/80) notes, production of onomatopoeia can serve as a form of "linguistic training" (p.27), through which "the phonemic system is...enriched" (p.27). Results from this thesis can now provide evidence in support of Jakobson's claims, showing that onomatopoeia do indeed facilitate lexical development through the stabilisation of the phonological system as a result of ample vocal practice.

8.3 Salience in the input

In this thesis we have observed how the 'special status' of OWs prompts the use of salient prosodic features in caregivers' production of these forms. That is, their lexical function as imitations of natural sounds leads to the use of pitch, rhythm and consonant/vowel lengthening to realistically represent the given sound. Inevitably this leads to these forms being particularly salient in the input, most notably in relation to their equivalent CW forms. While we considered OWs in terms of the typically-reported characteristics of IDS in Chapter 6, including higher pitch, pitch modulations and use of pauses (Cristia, 2013; Soderstrom, 2007), the added 'special effect' features observed in Chapter 7 demonstrate how these forms go beyond the typical prosodic cues of input speech. Here, infants are presented with an especially affective speech style, which combines even further-exaggerated IDS features with articulatory modifications such as growling, nasalised segments and varied phonation-types.

There is ample research to show that infants are sensitive to even the most subtle prosodic cues from a very young age: White and colleagues (2014) showed that British infants as young as five months attend to utterance-final durational cues in prosodic structure, while German (but not French) six-month-olds have been shown to prefer a trochaic stress pattern, as is typical of their ambient language (Höhle et al., 2009). Johnson and Jusczyk (2001) found that 8-month-old American infants were more sensitive to stress cues in input speech than statistical regularities, while a study by Mattys and colleagues (1999) showed that, by nine months of age, English-learning infants are sensitive to phonotactic, as well as prosodic, cues in the input. The authors propose that prosodic and phonotactic regularities in the input allow infants to infer word boundary positions, thus aiding the segmentation of running speech.

Infants' awareness of the prosodic detail of the input at such a young age suggests that prosodic features may be especially important in early language perception. This is in line with the 'prosodic bootstrapping hypothesis' (Gleitman & Wanner, 1982; but see Fernald & McRoberts, 2009, for a critical review), and it is suggested that the affective use of prosody in IDS may also contribute to this bootstrapping effect: Morgan and Demuth (1996) state that "children's perceptual analyses of speech must come before semantic analyses of representations of input utterances" (p.16). Indeed, numerous studies have shown how infants attend preferentially to speech presented in the typical 'motherese' style (Fernald & Kuhl, 1987; Fernald et al., 1989), which suggests that the features used in caregivers' production of OWs would also hold infants' attention more successfully than their less-salient CW equivalents. This was shown in Chapter 4, where infants attended longer to the target image upon hearing stimuli with a higher mean pitch. Duration was also thought to have an influence on infants' responses here, but the results were potentially distorted by a ceiling effect.

Findings from Chapter 6 also revealed how the use of pauses and isolated words are prevalent in mothers' production of OWs. This corresponds to the literature showing that these features are important in the early input (Brent & Siskind, 2001; Keren-Portnoy et al, in press; Ninio, 1993), as they enable infants to segment individual words from the speech stream. Further to this, the considerable amount of reduplication that has been identified in caregivers' production of OWs may also provide a distinct segmentation advantage for these forms. We can consider these findings alongside Gervain and colleagues' (2008) results showing an advantage for repeated syllable sequences in early perception. Furthermore, when combined with Ambridge et al.'s (2015; see also Monaghan & Christiansen, 2010) discussion on 'frequency effects' in language acquisition we find clear evidence to show that the use of repetition and reduplication may support a learning advantage for OWs in early language development.

In all, the findings from this thesis allow us to propose that in many cases, OWs are presented to infants in a particularly salient speech register. This incorporates the typical features of IDS with more idiosyncratic sound effects to create an especially affective speech style, which is exclusive to the production of these forms. Evidence from the literature on both IDS and infant sensitivity to prosody suggests that this provides a distinct learning advantage for OWs over the rest of the speech stream; infants are known to draw upon these features in word discrimination and segmentation, and also to attend preferentially to prosodically salient forms. Indeed, the prosody observed in mothers' production of OWs combines attention-grabbing features with segmentation prompts, thus facilitating the learning of OWs above their equivalent CW forms, and no doubt leading to the prominence of these forms in the early lexicon.

8.4 Wildness and communication

We have established that infants' early capacity to control features of pitch, rhythm and intensity may be highly relevant to their early production of OWs. Evidence from Chapter 7 suggests that this may be particularly important for early involvement in conversations, as infants can make referential use of wildness through the prosodic differentiation of protowords in the absence of the necessary phonological abilities. We have seen how the use of pitch in OW production can make these forms meaningful in early conversations, as the distinction between, for example, a high-low pitch modulation and a low-pitched growling effect may be the only distinguishing feature across two phonologically-identical forms.

Infants' ability to make use of wildness in this way is an important aspect of these findings, as lexically-speaking, OWs are unique in their prosodic characteristics: they are the only word forms which can be produced with extra-linguistic sound effects, but yet they are equally meaningful without any such effects. Furthermore, these sound effects are specific to any particular OW, and so their wild features do not vary too widely between different speakers and even different languages. This is supported by evidence from Chapter 4, where Swedish infants were able to recognise wild OWs produced in unfamiliar languages despite phonological differences between unfamiliar forms and the equivalent Swedish OWs, although results were distorted here by discrepancies in the data. Wild features provide a meaningful 'prosodic layer' on top of the typical phonological form of the OW, which infants can draw upon in both perception and production.

Wildness thus appears to facilitate engagement in early protoconversations, led by the caregiver who moderates early dialogues according to their infant's developing language abilities. Caregivers presented OWs in specific syntactic frames, which varied with their infants' developing ability to take part in turn-taking routines. In Chapter 6, when infants were too young to take part verbally in interactions with the caregiver, OWs were presented in isolation, and immediately after the corresponding CW. The pairing of OW and CW, as well as the use of pauses to reflect the rhythm and timing of interactions (Jaffe et al., 2001), is similar to the dialogues that we observed in Chapter 7, at which point the infants were old enough to take part. Indeed, we propose that these observations from Chapter 6 reflect an early example of turn-taking, as the mothers fulfilled both sides of the dialogue with the use of repetitive syntactic and prosodic features to present a 'call and response' style interaction.

Once infants are able to take part in turn-taking routines, as in Chapter 7, we see similar syntactic frames and consistent OW-CW pairings, but this time these are shared between infant and caregiver, as opposed to the caregiver taking on both 'turns'. Here the infant is continually prompted to produce specific words, many of which are OWs; the use of wildness allows the infants to produce a range of forms which may otherwise be beyond their articulatory capacity. Here we posit that the use of wildness facilitates the development of both phonological and conversational skills from a very young age, as OWs provide an extensive lexicon of material that can be used in turn-taking routines. Early on, the salient features of wild OWs draw infants' attention to the caregivers' production routines, while later this enables infants to produce forms which would otherwise be motorically out of reach.

8.5 OW vs. CW

Vygotsky's (1978) zone of proximal development (ZPD) was discussed in relation to the trends observed in turn-taking dialogues (Chapter 7), as it seems that OWs present caregivers with an opportunity to mediate their infants' learning through these early interactions. This highlights how the caregiver input may lead infants first to the acquisition of OW forms, and then to the acquisition of the equivalent CWs as phonological capacity develops. Results from Chapters 2 and 3 can be considered in light of these findings, as we observed this same trajectory in infants' acquisition of OWs and CWs: in almost all cases, the OW was acquired first.

In Chapter 7 we observed how the caregivers' input mediated the infants' OW production as phonological capacity developed. Over time, the caregivers produced more CWs than OW forms, as the infants' articulatory skills came together to bring about independent OW production as opposed to simple repetitions of the caregivers' own speech. Elsen (1994) describes her own efforts to encourage Annalena to move from OW to CW production; she continually prompted the infant to produce the CW *Hund* 'dog' at 0;11, at which point the infant was producing the OW *wauwau* 'woof woof' in every possible context. Here we observe an interaction between the infants' output and phonological capacity at any point in time, and the caregivers' mediation of this in relation to both the CW and the ZPD: once the infant has demonstrated the ability to produce an OW, the obvious next step in the trajectory of lexical development is the CW – the 'adult' equivalent of the infant's form. In the case of Annalena, Elsen (1994) reports a delay of three months between her initial attempts to prompt the adult form and Annalena's first production of *Hund*; this demonstrates how acquisition of the CW is driven primarily by the infant, mediated by a combination of caregiver input and developing phonological capacity. Indeed, we observe this in Chapters 2 and 3, when infants' outputs were clearly driven by their phonological abilities. Only when the necessary phonological changes had taken place did infants begin to move towards CW forms, and evidence from Chapter 7 shows how this is guided by the caregivers' input.

8.6 OWs and symbol formation

The combined findings from the six studies in this thesis have led us to propose that infants' abundant production of OWs in early lexical development is driven by influence from the input alongside the infant's phonological capacity at any point in time. We have found clear evidence to show that Werner and Kaplan's (1963) symbol formation hypothesis is not sufficient in explaining infants' production of OWs; importantly, we have shown a direct correlation between the infants' experience of OWs in the input and the eventual production of these forms in the output. According to these results, if a caregiver does not produce OWs in IDS, their infant will not go on to acquire these forms in the early lexicon.

This point is in fundamental conflict with Werner and Kaplan's position, since they propose that "at…early levels of representation, the vehicle is produced…as a mimetic facsimile of the referent" (1963, p.47). That is, from their perspective the acquisition of onomatopoeia is driven by infants' tendency to imitate sounds – including the imitation of sounds in their environment as well as that of the adult's production of onomatopoeia to which they are drawn – as a first step towards understanding the arbitrary symbol-referent relationship. Onomatopoeia, both environmental and lexical, are less differentiated from their referents, and thus they allow infants to mediate the difficult concept of arbitrariness in language. This is overcome through 'increasing differentiation', as infants slowly move towards the more conventional and arbitrary form.

Werner and Kaplan describe infants' production of 'transitional forms' (p.107) as they come to terms with the distance between symbol and referent. However, the analyses here have provided us with an alternative explanation for infants' use of OW-CW combinations such as *choo choo train* and *baa sheep* (examples taken from Stern, 1928, cited in Werner & Kaplan, 1963, p.108). Namely, in the majority of cases in Chapters 6 and 7, OWs were presented to the infant in combination with the equivalent CW, which often occurred immediately next to the OW. It is therefore unsurprising that, as infants' articulatory capacities develop, their ability to produce both OW and CW forms results in the combination of the two, as experienced in the input. In addition, it seems that in reality these combinations, or 'transitional forms', are the exception in the infant output rather than the rule, as only a couple of examples could be found across the nine datasets analysed in Chapter 2: [aviʃkabu:] (*kráva-bu* 'cow-moo', P, 1;6) and [m:kɯç] (*muh-Kuh* 'moo-cow', Annalena, 1;5).

The data analysed in this thesis provides some evidence towards infants' imitation of sounds from the environment, which could be interpreted as fitting with Werner and Kaplan's overall hypothesis regarding onomatopoeia. With regard to innovative OWs, we have found no clear evidence to suggest that it is infants' experience of these forms in the input that drives their early acquisition, and in the case of Annalena at least, it is confirmed that they were created by the infant (Elsen, 1991). Our findings lead us to return once again to Jakoson's propositions regarding the importance of onomatopoeia in early production. Jakobson (1941/80) discusses onomatopoeia with regard to infants' "desire to communicate..[and] the ability to communicate something" (p.24): the use of rudimentary vocal gestures to imitate sounds in the environment allows infants to articulate meaningful units, which we now know to be understood by the caregiver. Indeed, we have already discussed infants' ability to imitate pitch features from a very young age, and so it could be proposed that this tendency might also extend to sounds in the environment. Kessen and colleagues' (1979) findings support this proposal, since the infants in their study were found to imitate pitch tones whether these were sung by the mothers or played on a pitchpipe. Here we see that infants mimic not only the pitch of speech but also the musical modulation of pitch, and this may also extend to the imitation of salient pitch changes in the environment.

Infants' use of extra-phonetic features such as grunts and lip-smacking, as observed in vocal play, may also be drawn upon here, as Annalena is reported to produce features such as a bilabial trill, an interdental fricative and numerous vocalic sounds in the production of her innovative OWs. Before word production is properly established, the use of these vocal features – which approximate environmental sounds more closely than they do words – may be drawn upon simply to maximise output production despite a limited vocal capacity. This may function as a form of 'articulatory filter', as in the early stages of production infants are able to match environmental sounds with their rudimentary vocal features more closely than many of the words heard in caregiver speech. Infants' imitation of sounds from the environment is reported in many accounts of early word production (see Jakobson, 1941/80 and Werner and Kaplan, 1963, for numerous examples from the literature), and while these forms did not constitute the majority of OWs in our analyses, such onomatopoeic creations highlight just one aspect of OWs that may govern their presence in early infant speech. To return to Jakobson's point, these innovative OWs speak to infants' wish for "the expressive value of the extraordinary, rather than the desire for faithful sound imitation" (Jakobson, 1941/80, p.26). That is, innovative OWs offer the opportunity to produce words when almost all phonological resources are restricted.

8.7 OWs and iconicity

Finally, we return to the question of iconicity, and the extent to which this is relevant in OW acquisition. Recent research has tended to discuss OWs as a manifestation of sound symbolism: Nygaard and colleagues (2009) claim that onomatopoeia are "one of the most obvious examples" (p.181) of non-arbitrary sound-symbol correspondences, while Imai and Kita (2014) discuss onomatopoeia as part of sound symbolism alongside Japanese mimetics. Findings in this thesis lead us to propose that this is not an appropriate comparison, as many questions have been raised here regarding the nature of iconicity in OWs.

A large body of recent research has identified an early sensitivity to sound symbolic correspondences (e.g. Imai et al., 2015; Ozturk et al., 2013; Peña et al., 2011), which Ozturk and colleagues (2013) have suggested may be intrinsic to the human sensory system. Accounts of sound symbolism consistently propose that multi-modal synesthetic mappings may facilitate the learning and retrieval of form-meaning correspondences, as discussed most comprehensively in Ramachandran and Hubbard's (2001) account. However, this conclusion does not hold when onomatopoeia are considered within this framework: while infants may show symbol-object sensitivity when presented with forms such as *bouba* and *kiki*, authors consistently refer to the phonetic properties of these non-words and the correspondences that they may share with their corresponding round or spiky referents (e.g. D'Onofrio, 2014). This relates to

Ohala's (1984) 'frequency code', and the idea that specific phonetic properties can be identified symbolically in a referent. This becomes problematic when we consider onomatopoeia from this perspective, since it is often reported that infants produce OWs in reference to the relevant animal or vehicle, such as Annalena's production of *wauwau* in reference to dogs and dog-related objects (Elsen, 1991). In this case, it could be claimed that the referent DOG is further-removed from the OW *woof woof* than from the conventional word *dog*, since the word *woof woof* is iconic in relation to the sound that a dog makes and not actually the dog itself. From this perspective it seems that, symbolically speaking, this could even be a more difficult word-referent association to make.

Importantly, while infants may show sensitivity to congruence between non-words and symbolically matching shapes, this has no real bearing on word learning from experience of language. Indeed, the input exposes infants to countless word-object mappings, all of which are to a certain extent multi-modal owing to the simultaneous integration of articulatory and auditory, as well as referential, parameters (Westermann & Miranda, 2004). Moreover, in most cases sound symbolic congruence is not relevant to infants' acquisition of a word form. While the bouba-kiki effect may provide a fascinating insight into human language processing, in reality this is not an appropriate representation of our experience of speech in real-time.

We also discussed iconicity in relation to wildness, as Rhodes (1994) appears to suggest that it is wildness that makes an OW iconic, rather than any phonemic proximity between an OW and its equivalent real-world sound. The notion of iconicity in OWs becomes clearer when considering the relationship between the wild OW form and the sound that it seeks to approximate, but in fact this complicates the question of iconicity in early OW production. As observed in Chapter 7, tame OWs appear to be more commonly produced in caregiver-infant interactions than their wild equivalents, and again we must return to the question of the referent here, which is often a type of animal or vehicle rather than the real-world sound itself. Furthermore, alongside the potential advantage for iconicity in these forms, we also have confounding evidence which shows that they are more salient than their equivalent CWs in caregiver speech (Chapter 6). Wild OWs appear to provide a perceptual advantage owing to the extra prosodic salience of the wild features, and so it is difficult to separate the iconicity debate from the potential for a perceptual advantage in wild OWs. Even so, the hypothesis proposing that onomatopoeia (whether wild or tame) may have a learning advantage in early language development remains theoretical, while the advantage for salient prosodic features in IDS is covered extensively in the literature.

With these points in mind, we are led to question whether the iconicity argument is still a valid one in infant language development. Specific to this thesis, it does not seem appropriate to consider onomatopoeia in the same discussion as sound symbolism; the bouba-kiki effect and mimetics both apply Ohala's (1984) 'frequency code' to some extent in their sound-meaning relationships, which is surely what defines the multimodal correspondences that are proposed to be advantageous in the acquisition of sound symbolic forms. Onomatopoeia, on the other hand, do not represent a direct symbol-object relationship, and the frequency code does not apply here; instead, the phonemic properties of the word are determined by the human approximation of the real-world sound, which has become conventionalised over centuries of use (de Saussure, 1915/62). More generally, the sound symbolism literature tests participants almost exclusively on their perception of non-words in relation to indiscriminate round or angular shapes (though cf. D'Onofrio, 2014, for an analysis using real-wold objects and non-words). This makes it impossible to consider the relevance of sound symbolism in reality, since we cannot relate participants' experience of a word or object in an experimental setting to their perception and experience of it in the real world. Furthermore, while participants may be better able to learn symbolic sound-meaning correspondences in a novel-word learning task (Imai et al., 2008; Kantartzis et al., 2011) this does not necessarily reflect the reality of language learning, which always involves numerous contexts and modalities that cannot be controlled for in an experiment. Monaghan and colleagues (2014) attempt to address this issue using a corpus-analysis, but the omission of data from infants under two years sidesteps the most relevant period of infant language acquisition.

8.8 Limitations

The analyses in this thesis have allowed us to identify a role for onomatopoeia in infant language development. However, while the findings are multifaceted, still it has not been possible to fully answer the main question posited in the literature on this topic: do the iconic properties of onomatopoeia serve to facilitate the learning of these words in early language development? While it is tempting to suggest that the collective results from the six separate analyses provide a clear picture of how onomatopoeia function in early language learning, in fact none of these results are sufficient to make any outright claims against the iconicity hypothesis. In part, methodological issues led us to question results that might have otherwise addressed this point, but still important theoretical questions surrounding iconicity remain which are difficult to fully unpick.

Methodological issues in the first eye-tracking study (Chapter 4) distorted evidence that might have otherwise shown clearer results regarding infants' perception of wild and tame OWs. The second eye-tracking study (Chapter 5) accounted for some of these issues, but here it was unclear whether the lack of difference across conditions reflected the absence of any iconic advantage, or if the participants – most of them still unable to speak – were simply not responding to the task in the expected manner. A breakdown of results in accordance with the accompanying CDI results suggests that the former may be the case, as the words reported to be known by the infants generated the fastest responses across targets, but still this was not convincing enough to allow us to draw any strong conclusions.

It is also apparent that the specific nature of iconicity is very difficult to interrogate, as this concept is multifaceted and applicable to a number of differing aspects of linguistics. The consideration of onomatopoeia and sound symbolism as one and the same is problematic, as there is a notable discrepancy between auditory-visual correspondences in one respect and natural sound-vocal correspondences in the other: the only clear connection between these two features of language is the ability to generate iconic representations from them. In this respect, providing evidence which allows us to consider both aspects of the literature has not been possible, since infants' perception of onomatopoeic forms cannot plausibly be related to findings from the sound symbolism literature. Thus, while the findings from this thesis allow us to make claims regarding infants' early production of onomatopoeia, this cannot be extended to a thorough discussion of sound symbolism.

Determining the exact nature of iconicity in OWs has also presented issues, since it is unclear to what extent onomatopoeia truly represent their associated sound. In her extensive thesis on this topic, Masuda (2002) provides an acoustic-phonetic analysis of onomatopoeia in relation to their natural sounds. Here she shows that the formant dynamics of the onomatopoeic word parallel those of the equivalent sound, thus demonstrating that these forms do bear some resemblance to their real-world referent. However, this brings us back to the fundamental question of infants' experience of real-world sounds, since this is required if the infant is to draw upon this potential iconicity in onomatopoeia. In many cases it can be assumed that an infants' production of OWs is largely dependent on their caregivers' use of these forms: examples (5), (11) and (25) in Chapter 7 show the use of OWs such as a lion's roar, a bear growling and an owl hooting, all of which we can safely conclude have been learned through the input, and not through the infant's own experience of these sounds in the real world. However, this question is complicated further when we consider wildness as a feature of OWs, and how this may interact with or even determine their iconicity. Wildness inevitably increases the salience of OWs, but it could also render them more iconic. It is not possible to truly separate these two issues in cases such as *woof woof*, where infants are assumed to have experience of both the real-world sound and both its wild and tame OW approximations. Consequently this leaves us with two unanswered questions that could not be addressed in this thesis: does experience of the real-world sound have any bearing on infants' learning of the equivalent OW, and does the use of wild features make the OW more iconic or simply more salient?

This brings us to infants' creation of onomatopoeia from natural sounds – 'innovative' OWs – which constitute a small but fascinating aspect of infant OW production. It has been proposed here that this is due to the accessible prosodic features that allow infants to create meaningful units from these sounds, but this hypothesis is only speculative, and cannot be supported by any empirical evidence owing to the small number of these forms in the comparatively large datasets analysed in this thesis. Accounts of onomatopoeia in the literature discuss this phenomenon in detail (Jakobson, 1941/80; Werner & Kaplan, 1963), using single examples from a number of different infants, and Werner and Kaplan (1963) even suggest that these forms constitute the very beginnings of infant speech. However, the authors do not present a holistic picture of these infants' early outputs, and so it is unclear whether these forms constitute their very first words, or whether they represent only a fraction of an otherwise conventional lexicon. As observed in Chapters 2 and 3, a full picture of an infant's early lexicon is necessary in the consideration of their OW production.

Finally, perhaps the main limitation of this thesis as a contribution to linguistic research is the narrow spectrum of languages considered overall. While attempts have been made throughout to incorporate as many languages as possible, limitations on both the author's part and the availability of data (as well as the more practical limits imposed by this thesis) have determined the extent to which different languages can be considered. A total of eight different languages have been included in these analyses, incorporating a range of prosodic and phonological features which make for an interesting comparison. Despite this, we have noted observations in the literature which cannot be supported by findings here: Kunnari's (2002) account of Finnish syllabification, for example, found that infants produced a high number of disyllables in their early words to match the typical syllable structure of the ambient language. However, many OWs are monosyllabic in Finnish, and these were produced accurately by the infants, thus contrasting with the trends observed elsewhere in their output. Owing to this discrepancy, Kunnari analysed OWs separately from the rest of the data, as these forms were otherwise found to distort her findings. Similarly, Vihman (1976) presents data of her daughter's first fifty words, where we see how reduplicated OWs in the target form are truncated into single syllables (e.g. *cutcutcut* 'cluck cluck' [kɔ], *bow-wow* [au]); here, the apparent ubiquity of reduplication in early development is explicitly avoided, and instead the infant adapts the forms to match her own output preferences. These examples demonstrate the importance of a broad cross-linguistic analysis wherever possible in child language research; we may find that OWs play different roles in different languages, especially when their simple structures do not match the typical features of the ambient language.

8.9 Implications and directions for future research

The findings in this thesis allow us to view OWs from a new perspective, showing them to fit within the general systems that are often observed in infant production. This leads us to the conclusion that they should not be omitted from the study of infant language; onomatopoeia should be considered as part of the general trajectory to full language use. While Kunnari's (2002) observations contradict this suggestion, the role of these forms in an infants' experience of language – in both perception and production – cannot and should not be overlooked. These findings highlight how OWs enable production in the absence of phonological capacities, and a role has been established for these forms in the promotion of production in spite of these segmental limitations. With this in mind, to overlook OW production in an infant's early output is to exclude an essential mechanism in the dynamic trajectory of development – that is, production practice, in whatever form it may take.

From a theoretical perspective, it has been established that sound symbolism and onomatopoeia do not overlap in a way which allows them to be considered within the same domain. With this in mind, we propose that the iconicity literature avoids discussing onomatopoeia and sound symbolism as one and the same. The question of iconicity in onomatopoeia remains relevant, but the boundaries between these two aspects of language do not overlap, and in fact they appear to lie much further apart than is generally implied. With this in mind, any advantages found in sound symbolic correspondences should not be automatically assumed to also apply to onomatopoeia.

To return to the debate between Cratylus and Hermogenes (Jowett, 1953) introduced in Chapter 1, it seems that the question of arbitrariness in language is not as dichotomous as this dispute would suggest. Furthermore, Socrates' renunciation of the formation of words through the imitation of natural sounds as "wild and ridiculous" (426b) has been refuted empirically by the data observed here. With this in mind, we must perhaps begin to consider the nature of onomatopoeia in more detail, relating to the question of iconicity and precisely how these forms fit within the definition of 'iconic'. Infants' creation of onomatopoeic imitations, as well as the use of wildness in OW production, have led us to consider the different possibilities for iconicity in onomatopoeic depiction. When it comes to infant language at least, it seems that these forms present differing aspects of iconicity depending on the infants' experience of both the natural sound in question (recall the infant mentioned in Chapter 4, who had a number of cats at home and produced a high-pitched sound in reference to cats) and their experience of the lexicalised OW in the input. This opens up many further lines of enquiry, and while gathering sufficient data to carry out an analysis of some of these more specific aspects of OWs would no doubt be challenging, the question of how the use of OWs differs according to learned versus created forms would address some of the more established theoretical perspectives found in the literature, including both Werner and Kaplan (1963) and Jakobson's (1941/80) accounts.

Our questions regarding infants' responses to OW versus CW forms have still not been sufficiently analysed, and perhaps a replication of the eye-tracking experiment in Chapter 5 with slightly older infants would reveal clearer findings. However, a different methodological approach could also be applied here – namely the use of EEG techniques – which would enable the testing of pre-linguistic infants and, moreover, would confirm definitively whether or not OWs and CWs are processed equally in sound-object mapping tasks. A further possibility for any replication of this study would be a comparison of infants' responses to OWs against the equivalent real-world sounds, in order to establish truly how integrated – and thus iconic – these forms are in early language learning.

8.10 Conclusion

This thesis has provided an exploratory analysis of onomatopoeia in infant language development. Overall, we have shown that onomatopoeia do provide a certain bootstrapping advantage in early language production, but contrary to Imai and Kita's (2014) sound symbolism bootstrapping hypothesis, this is not manifested in iconicity. Instead, we observed how the unique status of these forms offers infants a stepping stone into word-learning via advantages in perception, production and interaction. The key findings can be related to the three main questions that were raised in Chapter 1:

Onomatopoeia are more salient in the input. Firstly, these forms have been found to be more salient than their conventional counterparts in the early input. This promotes the acquisition of onomatopoeia over their conventional equivalents from the very outset of language learning, with prosodic advantages in terms of mean pitch, pitch range, duration, and isolation in the speech stream. Importantly, the use of extensive reduplication means that infants are exposed to OWs more frequently than other words, thus further supporting the early learning of these forms. We can also consider wildness from the perspective of perception, and how wild features may make OWs more memorable to infants – certainly OWs stand out from the speech stream with respect to both their prosodic and extra-phonetic features.

Onomatopoeia fit early production patterns. Common prosodic structures reported in accounts of phonological development typically involve reduplication, simple CV syllables and consonant harmony (Smith, 2004; Vihman, 2015). Across the data analysed here, onomatopoeia were found to fit these structures automatically, and thus it was proposed that these forms may be 'set up' for infants' early word production. Furthermore, we observed how OWs are a more suitable match to infants' early output patterns than their CW equivalents, and thus we observed the prioritisation of OW acquisition in early development. This was extended to segmental capabilities in Annalena's data (Chapter 3), whereby OWs were produced until she had mastered the phonemes and structures required to produce the CW forms.

Onomatopoeia are functional in early interactions. Over the course of development OWs played an important role in turn-taking routines, scaffolding proto-conversations between the infant and the caregiver through rehearsed interactions which developed over time in line with the infants' language abilities. An important role for wildness was also observed, as wild features made otherwise rudimentary vocal production meaningful and thus functional in early dialogues. Here we observed how the use of OWs in early interactions changed in line with the infants' developing language capacity, and this appeared to be mediated by the caregivers during turn-taking routines.

There is no doubt that OWs are peripheral to language, but nevertheless they have been found to fit within the framework of phonological development. They are produced in abundance by both infant and caregiver in situations of joint attention – these forms are thus relevant to the infant's early experiences, despite being rare in the adult language. Furthermore, while their marginal status appears to have led OWs to be overlooked in much of the developmental literature, this may be an important aspect of their role in language learning and should not be dismissed. We have already observed how their position at the peripheries of language allows infants more flexibility in the production of OWs through their loose phonological specification, but we must also consider their lexical position in terms of early word learning. Indeed, OWs offer a lexical alternative to many CW forms, and have been found to provide a simpler phonological structure for the infant to deal with, as well as the option for prosodic and extra-linguistic effects which render these forms both meaningful and memorable in early phonological development.

APPENDICES

APPENDIX I – ACCOMPANYING ETHICS DOCUMENTS

I.A Information sheet (English translation), eye-tracking experiment 1 (Ch. 4)

THE UNIVERSITY *of York*

DEPARTMENT OF
LANGUAGE AND
LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Enquiries cs524@york.ac.uk

INFORMATION SHEET

PLEASE KEEP THIS INFORMATION SHEET AND A SIGNED COPY OF THE CONSENT FORM FOR YOUR RECORDS

You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.

Title of study: Exploring the Role of Onomatopoeia in Phonological Development
Researcher: Catherine Smith

What is the research about?
This research looks at infants' responses to onomatopoeic forms produced in different ways in languages which are familiar and unfamiliar to the infant.

Why is the research being carried out?
This experiment forms part of a PhD project investigating the role of onomatopoeic word forms in infant's early word use. It will provide evidence towards why onomatopoeia is so common in words spoken by young infants.

Who is carrying out the research?
The research is being carried out by Catherine Smith of the Department of Language and Linguistic Science at the University of York.

Who can participate?
As this study is looking at infant's word production, only infants of a certain age can take part.
Your infant can participate if:

(i) He/she is aged between 14-16 months of age.
(ii) He/she is growing up within a monolingual language setting, learning only Swedish
(iii) He/she is not exposed to any language other than Swedish at home or in a caregiving situation.

What does the study involve?
We will ask you to fill in a brief questionnaire about the words that your infant produces and understands. We will then ask you to hold your infant on your knee while he/she takes part in an eye-tracking experiment. You will be provided with headphones to ensure that you cannot hear the stimuli that your infant can hear – this is to ensure that there is no possibility that your infant's responses have been influenced. Your infant will be presented with a series of 24 displays, each showing photos of two animals. A corresponding sound will be heard over the audio after a couple of seconds, and your infant's eye movements will be recorded.

Does my infant have to take part?
Your infant does not have to take part in the study. If you do decide that he/she will take part you will be given this information sheet to keep and will be asked to sign two copies of the consent form (one copy is for you to keep). If you decide to allow your infant to take part you will still be free to withdraw him/her at any time without giving a reason. If you decide to withdraw your infant from the study, we will destroy your data and will not use it in any way.

What are the possible risks of taking part?
The study poses no foreseeable risk to you or your infant. If your infant is distressed by the situation in any way then we can pause or stop the experiment as necessary.

Are there any benefits to participating?
You will be participating in important new piece of linguistic research that may help us to determine how infants learn language.

What kind of information do I have to give?
We will ask you to give some simple information about your family situation and your infant: the educational background of adults living in your home, the place where the adults living in your home grew up, your infant's primary caregiver, your infant's age, and whether he/she has suffered any abnormalities in his/her development.

What will happen to the data I provide?
The data you provide will be kept safely on university computers and will be backed up. It will be analysed to find out how infants respond to onomatopoeic words. The results will be reported in presentations and one or more academic papers. The data may also be kept after the duration of the current project, to be used in future research on language development.

What about confidentiality?
Your identity will be kept strictly confidential. No real names will be used in any presentations or publications. All data will be stored electronically under a code name. The only place your name (and, optionally, your email address) will be recorded will be the consent form, which will be stored securely.

Will I know the results?
We will not be able to give you feedback on your individual results.

*This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions regarding this, you can contact the head of the Ethics Committee, Tamar Keren-Portnoy (**email:** tamar.keren-portnoy@york.ac.uk; **Tel:** (+44) 01904 323614).*

If you have further questions regarding this study, please feel free to contact:

Catherine Smith
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
email: cs524@york.ac.uk

I.B Consent form (English translation), eye-tracking experiment 1 (Ch. 4)

Exploring the Role of Onomatopoeia in Phonological Development

Lead researcher: Catherine Smith

Consent form

This form is for you to state whether or not you agree for your infant to take part in the study. Please read and answer every question. If there is anything you do not understand, or if you want more information, please ask the researcher.

Have you read and understood the information leaflet about the study?	Yes ❒ No ❒
Have you had an opportunity to ask questions about the study and have these been answered satisfactorily?	Yes ❒ No ❒
Do you understand that the information that you and your infant provide will be held in confidence by the research team, and names or identifying information about you and your infant will not be mentioned in any publication?	Yes ❒ No ❒
Do you understand that you may withdraw your infant from the study at any time before the end of the data collection session without giving any reason, and that in such a case all his/her data will be destroyed?	Yes ❒ No ❒
Do you understand that the information you provide may be kept after the duration of the current project, to be used in future research on language?	Yes ❒ No ❒
Do you give consent for your infant to take part in the study?	Yes ❒ No ❒
Do you agree to the researcher's keeping your contact details after the end of the current project, in order that s/he may contact you in the future about possible participation in other studies? *(You may take part in the study without agreeing to this).*	Yes ❒ No ❒

Your name (in BLOCK letters): ______________________________

Your signature: ______________________________

Researcher's name: ______________________________

Date: ______________________________

One copy to be retained by the researcher, one copy to be kept by the participant.

I.C Information sheet, eye-tracking experiment 2 (Ch. 5)

THE UNIVERSITY of York

DEPARTMENT OF
LANGUAGE AND
LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Enquiries catherine.laing@york.ac.uk

INFORMATION SHEET

PLEASE KEEP THIS INFORMATION SHEET AND A SIGNED COPY OF THE CONSENT FORM FOR YOUR RECORDS

You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.

Title of study: Exploring the Role of Onomatopoeia in Phonological Development
Researcher: Catherine Laing

What is the research about?
This research looks at infants' responses to words which are common in 'babytalk'.

Why is the research being carried out?
This experiment forms part of a PhD project investigating the role of onomatopoeic word forms in infant's early word use. It will provide evidence towards why onomatopoeia are so common among words spoken by young infants.

Who is carrying out the research?
The research is being carried out by Catherine Laing of the Department of Language and Linguistic Science at the University of York.

Who can participate?
As this study is looking at infants' understanding of babytalk words, only infants of a certain age can take part.

Your infant can participate if:

(i) He/she is under 12 months of age.
(ii) He/she is learning English as a main language, and British English is the main language spoken at home.
(iii) He/she does not have any speech, sight or hearing impairments.
(iv) He/she does not have any developmental disorders.
(v) He/she was not delivered prematurely at birth.

1

What does the study involve?
You will complete a brief questionnaire about the words that your infant understands, as well as the words that he/she can produce if already able to talk. Your infant will then take part in an eye-tracking experiment, while you hold him/her on your knee. We will provide you with headphones and some blacked-out sunglasses – this will ensure that you cannot hear or see the stimuli, and that the eye-tracker won't pick up any of your eye movements by mistake. Your infant will see 28 displays, each showing photos of two objects, while we play a corresponding sound over the audio. We will then record your infant's eye movements. The experiment will take under 10 minutes to complete.

Does my infant have to take part?
Your infant does not have to take part in the study. If you do decide that he/she will take part you will be given this information sheet to keep and will be asked to sign two copies of the consent form (one copy is for you to keep). If you decide to allow your infant to take part you will still be free to withdraw him/her at any time during the procedure without giving a reason. If you decide to withdraw your infant from the study, we will destroy your infant's data and will not use it in any way.

What are the possible risks of taking part?
The study poses no foreseeable risk to you or your infant. If your infant is distressed by the situation in any way then we can pause or stop the experiment as necessary.

Are there any benefits to participating?
You will be participating in an important new piece of linguistic research that may help us to determine how infants learn language. Your infant will also be rewarded with a t-shirt as a thank you for participating in the study. He/she will receive this whether or not you are able to complete the experiment.

What will happen to the data I provide?
The data you provide will be kept safely on computers at the University of York and will be backed up. It will be analysed to find out how infants respond to babytalk words. The results will be reported in conference presentations and one or more academic papers, as well as a PhD thesis.

What about confidentiality?
Your identity will be kept strictly confidential. No real names will be used in any presentations or publications. All data will be stored electronically under a code name. The only place your name (and, optionally, your email address) will be recorded will be the consent form, which will be stored securely.

Will I know the results?
You have the option of receiving a summary of the results once the experiment is completed. Please indicate on the consent form if you would like to receive this.

*This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions regarding this, you can contact the head of the Ethics Committee, Dominic Watt (**email:** dominic.watt@york.ac.uk; **Tel:** (+44) 01904 322671).*

If you have further questions regarding this study, please feel free to contact:

Catherine Laing
PhD Student
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
email: catherine.laing@york.ac.uk

Professor Marilyn Vihman
PhD Supervisor
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
email: marilyn.vihman@york.ac.uk

Exploring the Role of Onomatopoeia in Phonological Development

Lead researcher: Catherine Laing

Consent form

This form is for you to state whether or not you agree for your infant to take part in the study. Please read and answer every question. If there is anything you do not understand, or if you want more information, please ask the researcher.

Have you read and understood the information leaflet about the study?	Yes ❐ No ❐
Have you had an opportunity to ask questions about the study and have these been answered satisfactorily?	Yes ❐ No ❐
Do you understand that the information that you and your infant provide will be held in confidence by the research team, and names or identifying information about you and your infant will not be mentioned in any publication?	Yes ❐ No ❐
Do you understand that you may withdraw your infant from the study at any time before the end of the data collection session without giving any reason, and that in such a case all his/her data will be destroyed?	Yes ❐ No ❐
Do you give consent for your infant to take part in the study?	Yes ❐ No ❐
Would you like to receive feedback on the results of this study once it is completed? If Yes, please provide an email address for us to send this feedback.	Yes ❐ No ❐

____

Your name (in BLOCK letters):

____

Your signature:

____

Researcher's name:

____

Researcher's signature

____

Date: ________

1

APPENDIX II – SIMILARITY SCORING

II.A Similarity scores across languages

Language	CAT		COW		DOG	
Swedish	/mjaːw/		/muː/		/vɔvɔv/	
Arabic	/miːaʊ/	.95	/mɵː/	.95	/hæwhæw/	.48
Chinese	/miːaʊ/	.95	/mɔː/	.78	/wʌŋwʌŋ/	.40
Urdu	/miːaʊ/	.95	/bʌː/	.73	/wɔwɔw/	.51
	DUCK		SHEEP		ROOSTER	
Swedish	/kvækvæk/		/bɛːbɛː/		/kukəliku/	
Arabic	/kwækwæk/	.74	/mɛːmɛː/	.96	/kukukuku/	.84
Chinese	/gægæ /	.57	/meːmeː/	.81	/ wə.wə /	.24
Urdu	/kwæ̃kwæ̃k/	.75	/meːmeː/	.81	/kukəɽũkəɽũ/	.63

Similarity scores are shown as decimals to the right of the transcription. Stimuli not used in the experiment are shown in grey.

II.B Similarity scoring calculations

Based on Mueller et al.'s (2003) PSIMETRICA model.

(1) ROOSTER: Urdu

L_F	ku	kə	li	Ø Ø	ku
L_U	ku	kə	ɽũ	kə	ɽũ
Phonemic similarity	1, 1	1, 1	.83, 60	0, 0	.38, .89
Syllabic similarity	1	1	.72	0	.64
Score	**0.67**				

(2) ROOSTER: Mandarin Chinese

L_F	ku	kə	li	ku
L_U	wə	Ø Ø	Ø Ø	wə
Phonemic similarity	.5, .79	0	0	.5, .79
Syllabic similarity	.64	0	0	.64
Score	**0.32**			

DUCK: Mandarin Chinese

L_F	kvæk	kvæk
L_U	gØæØ	gØæØ
Phonemic similarity	.93, 0, 1, .93, 0	.93, 0, 1, .93, 0
Syllabic similarity	.57	.57
Score	**0.57**	

List of abbreviations

ADS	adult-directed speech
f0	fundamental frequency
C	consonant
CH	consonant harmony
CW	conventional word (conventional equivalent of an onomatopoeic word)
ExPr	external prompting
F	familiar
Hz	Hertz (frequency)
IDS	infant-directed speech
Lab	labelling
L_F	familiar language
L_U	unfamiliar language
M	mean
ms	milliseconds
Overext.	Overextention
OW	onomatopoeic word
Redup.	reduplication
Rep	repetition
Resp	response
RW	regular word
s	seconds
SD	standard deviation
SpecPr	specific prompting
T	tame

U	unfamiliar
V	vowel
W	wild

REFERENCES

Abbott, D. (2004). *Animal sounds.* [Online]. [Accessed 9th March 2015]. Available from: http://www.eleceng.adelaide.edu.au/personal/dabbott/animal.html.

Acredolo, L. & Goodwyn, S. W. (1988). Symbolic gesturing in normal infants. *Child Development*, 59, 450-466.

Ambridge, B., Kidd, E., Rowland, C. F. & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42, 239-273.

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H. & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex*, 63, 196-205.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10, 48-53.

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. *Infancy*, 17, 126-140.

Barney, R. (2001). *Names and nature in Plato's Cratylus.* New York, NY: Routledge.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.

Bates, D. M., Maechler, M. & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes.* R package version 0.999999-0.

Baxter, T. M. S. (1992). *The Cratylus: Plato's critique of naming*. Leiden, Netherlands: E. J. Brill.

Bee, H. L., Barnard, K. E., Eyres, S. J., Gray, C. A., Hammond, M. A., Spietz, A. L., Snyder, C. & Clark, B. (1982). Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother-infant interaction. *Child Development*, 53, 1134-1156.

Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes*, 21, 2-24.

Bergelson, E. & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109, 3253-3258.

Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language*, 290-311.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, 5, 341- 345.

Boysson-Bardies, B. de & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, 297-319.

Boysson-Bardies, B. de, Sagart, L. & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11, 1-15.

Boysson-Bardies, B. de, Hallé, P., Sagart, L. & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.

Braunwald, S. R. (1976). Mother-child communication: The function of maternal language input. In R. von Raffler-Engel (ed.), *Child language – 1975* (pp. 28-50). Great Britain: International Linguistics Association.

Brent, M. R. & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition,* 81, B33-B44.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., Snow, K. L., Aslin, R. N., Haith, M. M., Wass, T. S. & Adler, S. A. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, 62, i-160.

Colman, A. M. (2009). *A dictionary of psychology*. Oxford: Oxford University Press.

Colombo, J., Richman, W. A., Shaddy, D. J., Follmer Greenhoot, A. & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, 72, 1605-1616.

Compton, A. J. & Streeter, M. (1977). *Child phonology: Data collection and preliminary analyses. Papers and reports on child language development*. ERIC Clearinghouse.

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass*, 7, 157-170.

Cristia, A. (2015, June). *Language acquisition: From the lab to the wild*. Plenary talk presented at the 2nd Workshop on Infant Language Development (WILD), Stockholm, Sweden.

Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology*, 52, 259-268.

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.

DePaolis, R. A., Keren-Portnoy, T. & Vihman, M. M. (2010). Exploring cultural impact on segmentation and first word recognition: *ESRC End of Award Report*, RES-000-22-3331. Swindon: ESRC.

DePaolis, R. A., Vihman, M. M. & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behaviour and Development*, 36, 642-649.

Deuchar, M. & Quay, S. (2000). *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.

D'Onofrio, A. (2014). Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm. *Language and Speech*, 57, 367-393.

Dromi, E. (1987). *Early lexical development*. San Diego, CA: Singular Publishing Group Inc.

Echols, C. H. & Newport, E. L. (1992). The role of stress and position in determining first words. *Language Acquisition*, 2, 189-220.

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.

Elsen, H. (1996). Two routes to language: stylistic variation in one child. *First Language*, 16, 141-158

Endress, A. D., Dehaene-Lambertz, G. & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, 105, 577-614.

Fantz, R. L. 1964. Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, 146, 668-670.

Farrar, F. W. (1883). *Language and languages*. London: Longmans, Green.

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.

Ferguson, C. A. (1964). Babytalk in six languages. *American Anthropologist,* 66, 103-114.

Ferguson, C. A. (1983). Reduplication in child phonology. *Journal of Child Language,* 10, 239-243.

Ferguson, C. A. & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 419-439.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development,* 60, 1497-1510.

Fernald, A. & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9, F33-F40.

Fernald, A. & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behaviour and Development*, 10, 279-293.

Fernald A. & McRoberts, G. (2009). Prosodic bootstrapping: a critical analysis of the argument and the evidence. In J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp.365-388). New York, NY: Psychology Press.

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

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A. & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9, 228-231.

Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-113.

Fernald A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B. & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477-501.

Fernald, A., Zangl, R., Portillo, A. L. & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. In I. A. Sekerina, E. M. Fernández & H. Clahsen (eds.) *Developmental Psycholinguistics: On-line methods in children's language processing* (pp.97-136). Amsterdam: John Benjamins.

Fikkert, P. & Levelt, C. (2008). How does place fall into place? The lexicon and emergent constraints in children's developing phonological grammar. In P. Avery, E. Dresher & K. Rice (eds.) *Contrast in phonology: theory, perception and acquisition* (pp. 231-270). Berlin: Mouton.

Fort, M., Weiß, A., Martin, A. & Peperkamp, S. (2013). Looking for the bouba-kiki effect in prelexical infants. *Proceedings of the 12th International Conference on Auditory–Visual Speech Processing, [Annecy, France, 29 August–1 September 2013],* 71-76.

Friedrich, M. (2008). Neurophysiological correlates of picture-word priming in one-year-olds. In A. D. Friederici & G. Thierry (eds.) *Early language development: Bridging brain and behaviour. TiLAR,* 5, 137-160.

Friedrich, M. & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding – weak retention. *Journal of Cognitive Neuroscience*, 23, 3228-3240.

Garrod, S. & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8, 8-11.

Gervain, J., Macagno, F., Cogoi, S. Peña, M. & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 14222-14227.

Gervain, J. & Werker, J. (2008). How Infant Speech Perception Contributes to Language Acquisition. *Language and Linguistics Compass*, 2, 1149-1170.

Gleitman, L. R. & Wanner, E. (1982). The logic of language acquisition. In E. Wanner & L. R. Gleitman (eds.) *Language acquisition: The state of the art* (pp.3-48). Cambridge, UK: Cambridge University Press.

Goldstein, M. H. & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19, 515-523.

Golinkoff, R. M. & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: implications for language acquisition. *Journal of Child Language*, 22, 703-726.

Goodman, M. B & Jusczyk, P. W. 2000. The onset of sensitivity to internal syllable structure. In M. B. Broe & J. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp.228-239). Cambridge, UK: Cambridge University Press.

Goodwyn, S. W., Acredolo, L. P. & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behaviour*, 24, 81-103.

Hallé, P. A. & Boysson-Bardies, B. de (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463-481.

Hamilton, A., Plunkett, K. & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, 27, 689-705.

Hargrave, A. C. & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15, 75-90.

Hayhoe, M. M. (2004). Advances in relating eye movements and cognition. *Infancy*, 6, 267-274.

Hinton, L., Nichols, J. & Ohala, J. J. (1994). Introduction: Sound symbolic processes. In L. Hinton, J. Nichols & J. Ohala, (eds.) *Sound symbolism* (pp.1-12). Cambridge, UK: Cambridge University Press.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, 32, 262-274.

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, 20130298.

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

Ingram, D. (1974). Phonological rules in young children. *Journal of Child Language*, 1, 49-64.

Ivanova, G. 2006. Sound-symbolic approach to Japanese mimetic words. *Toronto Working Papers in Linguistics*, 26, 103-114.

Iwasaki, N., Vinson, D. P. & Vigliocco, G. (2007). What do English speakers know about gera-gera and yota-yota? A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese-Language Education around the Globe*, 17, 53-78.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P. & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, 66, i-149.

Jakobson, R. (1980). *Child language, aphasia and phonological universals.* 3rd Edition, translated by Allan R. Keiler. The Hague, Netherlands: Mouton Publishers.

Jespersen, O. (1922). *Language: its nature, development and origin*. London, UK: Allen & Unwin.

Johnson, E. K. & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548-567.

Jowett, B. (1953). *The dialogues of Plato, vol. III.* 4th Edition. Oxford, UK: Clarendon Press.

Jusczyk, P. W., Hirsh-Pasek, K., Nelson, D. G. K., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.

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

Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, 39, 326-342.

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.

Kemler Nelson D. G., Hirsh-Pasek, K., Jusczyk, P. W. & Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16, 55-68.

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.

Keren-Portnoy, T., Vihman, M. M. & Lindop Fisher, R. (in press). Do infants learn from isolated words? An ecological study.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. & Williams, N. A. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language, and Hearing Research*, 53, 1280-1293.

Kern, S. (2010). Les premiers mots du jeune enfant français: analyse quantitative et qualitative du vocabulaire réceptif et productif des deux premières années de vie. *Rééducation Orthophonique*, 244, 149-165.

Kern, S. & Gayraud, F. (2010). *L'inventaire français du développement communicatif*. Grenoble, France: La Cigale.

Kessen, W., Levine, J., & Wendrich, K. A. (1979). The imitation of pitch in infants. *Infant Behaviour and Development*, 2, 93-99.

Kidd, C., Piantadosi, S.T. & Aslin, R.N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, 7, e36399.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics*, 35, 379-416.

Köhler, W. (1970). *Gestalt psychology: An introduction to new concepts in modern psychology* (9th ed.). New York, NY: New American Library.

Kovic, V., Plunkett, K. & Westermann, G. (2010). The shape of words in the brain. *Cognition*, 114, 19-28.

Kuhl, P. K. & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science,* 218, 1138-1141.

Kunnari, S. (2002). Word length in syllables: Evidence from early word production in Finnish. *First Language*, 22, 119-135.

Laeng, B., Sirois, S. & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7, 18-27.

Lee, S., Davis, B. L. & Macneilage, P. F. (2008). Segmental properties of input to infants: a study of Korean. *Journal of Child Language*, 35, 591-617.

Leopold, W. F. (1939). *Speech development of a bilingual child, 1: vocabulary growth in the first two years*. Evanston, IL: Northwestern University Press.

Levitt, A. G. (1993). The acquisition of prosody: Evidence from French-and English-learning infants. In *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 385-398). Dordrecht, Netherlands: Springer.

Levitt, A. G. & Aydelott Utman, J. G. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, 19, 19-49.

Lewis, M. M. (1936). *Infant Speech: A study of the beginnings of language*. New York, NY: Harcourt, Brace. Reprint edition, 1975. New York, NY: Arno Press.

Lieberth, A. K. & Gamble, M. E. B. (1991). The role of iconicity in sign language learning by hearing adults. *Journal of Communication Disorders*, 24, 89-99.

Lüke, C. (2015, March). I*conic gestures support novel word learning in children with specific language impairment*. Paper presented at the 37th Deutsche Gesellschaft für Sprachwissenschaft conference, Leipzig, Germany.

Lyons, J. (1968). *Introduction to theoretical linguistics*. Cambridge, UK: Cambridge University Press.

Macken, M. A. (1979). Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua*, 49, 11-49.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Magnus, M. (2001). *What's in a word? Studies in phonosemantics.* Unpublished PhD thesis. University of Trondheim, Norway.

Majorano, M., Vihman, M. M. & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development*, 10, 179-204.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, 92, 199-227.

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47, 241-275.

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

Mattys, S. L. & Jusczyk, P. W., Luce, P. A. & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-

494.

Maurer, D., Pathman, T. & Mondloch, C. J. (2006). The shape of boubas: sound–shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.

McCune, L. (2008). *How children learn to learn language*. Oxford, UK: Oxford University Press.

McCune, L. & Vihman, M. M. (2001). Early phonetic and lexical development. *Journal of Speech, Language and Hearing Research*, 44, 670-684.

McCune, L., Vihman, M. M., Roug-Hellichius, L., Delery, D. B. & Gogate, L. (1996). Grunt communication in human infants. *Journal of Comparative Psychology*, 110, 27-37.

McMurray, B., Kovack-Lesh, K. A., Goodwin, D. & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*, 129, 362-378.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143-178.

Meier, R. P., Mauk, C. E., Cheek, A. & Moreland, C. J. (2008). The form of children's early signs: iconic or motoric determinants? *Language Learning and Development*, 4, 63-98.

Meier, R. P. & Newport, E. L. (1990). Out of the hands of babes: On a possible sign advantage in language acquisition. *Language*, 66, 1-23.

Menn, L. (1971). Phonotactic rules in beginning speech: A study in the development of English discourse. *Lingua*, 26, 225-251.

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? The nature and sources of phonological primitives.*(Language Faculty and Beyond 6.) Amsterdam: John Benjamins.

Merin, N., Young, G. S., Ozonoff, S. & Rogers, S. J. (2007). Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-month-old infants at-risk for autism from comparison infants. *Journal of Autism and Developmental Disorders*, 37, 108-121.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91-117.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E. & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, 14, 115-128.

Monaghan, P. & Christiansen, M. H. (2010). Words in puddles of sound: modelling psycholinguistic effects in speech segmentation. *Journal of Child Language*, 37, 545-564.

Monaghan, P., Christiansen, M. H. & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology*, 140, 325-347.

Monaghan, P., Mattock, K. & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1152-1164.

Monaghan, P., Shillcock, R.C., Christiansen, M.H. & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B*, 369, 20130299.

Morgan, J. L. & Demuth, K. (1996). Signal to syntax: An overview. In J. L. Morgan & K. Demuth (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp.1-22). New York, NY: Psychology Press.

Morgenstern, A. & Parisse, C. (2007). Codage et interprétation du langage spontané d'enfants de 1 à 3 ans. *Corpus, 6,* 55-78.

Mueller, S. T., Seymour, T. L., Kieras, D. E. & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 29, 1353-1380.

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C. & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, 84, 2064-2078.

Namy, L. L., Campbell, A. L. & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, 5, 37-57.

Ninio, A. (1993). On the fringes of the system: Children's acquisition of syntactically isolated forms at the onset of speech. *First Language*, 13, 291-313.

Ninio, A., Snow, C. E., Pan, B. A. & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of Communication Disorders*, 27, 157-187.

Nygaard, L. C., Cook, A. E. & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, 112, 181-186.

Oakes, L. M. (2012). Advances in eye-tracking in infancy research. *Infancy*, 17, 1-8.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, 41, 1-16.

Orlansky, M. D. & Bonvillian, J. D. (1984). The role of iconicity in early sign language acquisition. *Journal of Speech and Hearing Disorders*, 49, 287-292.

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.

Papoušek, M., & Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language*, 9, 137-158.

Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition*, 6, 201- 253.

Peirce, J. W. (2007). PsychoPy - psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8-13.

Peña, M., Mehler, J. & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychological Science*, 22, 1419-1421.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46, 115-154.

Priddy Books. (2009a). *Home: Press-out flash card book*. London, UK: St. Martin's Press, LLC.

Priddy Books. (2009b). *Toys: Press-out flash card book*. London: St. Martin's Press, LLC.

Priestly, T. M. (1977). One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language*, 4, 45-65.

Psychology Software Tools, Inc. [*E-Prime 2.0*]. (2012). Retrieved from http://www.pstnet.com.

Ramachandran, V. S. & Hubbard, E. M. (2001). Synaesthesia – a window into perception, thought and language. *Journal of Consciousness Studies*, 8, 3-34.

R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

Rhodes, R. (1994). Aural images. In L. Hinton, J. Nichols & J. Ohala, (eds.) *Sound symbolism* (pp.276-292). Cambridge, UK: Cambridge University Press.

Rochat, P. (2001). Dialogical nature of cognition. *Monographs of the Society for Research in Child Development*, 66, 133-143.

Roder, B. J., Bushnell, E. W. & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, 1, 491-507.

Rondal, J. A. (1980). Fathers' and mothers' speech in early language evelopment. *Journal* *of Child Language*, 7, 353-369.

Rowe, M. L. & Goldin-Meadow, S. (2009). Early gesture *selectively* predicts later language learning. *Developmental Science*, 12, 182-187.

Sapir, E. (1970). *Language: An introduction to the study of speech*. London, UK: Rupert Hart-Davis.

Saussure, F. de (1962). *Cours de linguistique générale*. 5th Edition. Paris, France: Payot.

Savinainen-Makkonen, T. (2007). Geminate template: A model for first Finnish words. *First Language*, 27, 347-359.

Saxon, T. F. (1997). A longitudinal study of early mother-infant interaction and later language competence. *First Language*, 17, 271-281.

Seidl, A., Tincoff, R., Baker, C. & Cristia, A. (2015). Why the body comes first: Effects of experimenter touch on infants' word finding. *Developmental Science*, 18, 155-164.

Slobin, D. I., Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Pyers, J., Anthony, M., Biederman, Y. & Thumann, H. (2003). A cognitive/functional perspective on the acquisition of "classifiers". In K. Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Languages* (pp.271-296). Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, C. E. (2011). Variation and similarity in the phonological development of French dizygotic twins: Phonological bootstrapping towards segmental learning? *York Papers in Linguistics*, 2, 74-87.

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge, UK: Cambridge University Press.

Smith, N. V. (2004). Universal tendencies in the child's acquisition of phonology. In B. Lust & C. Foley (eds.), *First Language Acquisition: The essential readings* (pp.294-306). Oxford, UK: Blackwell Publishing.

Smith, N. A. & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13, 410-420.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501-532.

Soderstrom, M. & Wittebolle, K. (2013). When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. *PloS one*, 8, e80646.

Stark, R. E. (1978). Features of infant sounds: The emergence of cooing. *Journal of Child Language*, 5, 379-390.

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (eds.) *Child phonology 1: Production* (pp.43-72). New York, NY: Academic Press.

Stark, R. E. (1989). Temporal patterning of cry and non-cry sounds in the first eight months of life. *First Language*, 9, 107-136.

Stern, D. N., Spieker, S., Barnett R. K. & Mackain, K. (1983). The prosody of maternal speech: infant age and context related changes. *Journal of Child Language*, 10, 1-15.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, 38, 1-34.

Sundberg, U. (1998). *Mother tongue – phonetic aspects of infant-directed speech*. Stockholm, Sweden: Perilus XXI.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432-443.

Swingley, D. & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147-166.

Swingley, D. & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13, 480-484.

Swingley, D., Pinto, J. P. & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73-108.

Szreder, M. (2013). The acquisition of consonant clusters in Polish: a case study. In M. M. Vihman & T. Keren-Portnoy (eds.), *The Emergence of Phonology: Whole-word Approaches and Cross-Linguistic Evidence* (pp.343-361). Cambridge, UK: Cambridge University Press.

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

Taumoepeau, M. & Ruffman, T. (2008). Stepping stones to others' minds: Maternal talk relates to child mental state language and emotion understanding at 15, 24, and 33 months. *Child Development*, 79, 284-302.

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT press.

Tomasello, M. (2003). *Constructing a language: A usage-based approach to child language acquisition*. Boston, MA: Harvard University Press.

Turati, C., Di Giorgio, E., Bardi, L. & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants, and adults: Evidence from the composite face effect. *Child Development*, 81, 1894-1905.

Vihman, M. M. (1976). From pre-speech to speech: on early phonology. *Stanford Papers and Reports on Child Language Development*, 12, 230-244.

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.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological development: Models, research, implications* (pp. 393-422). Timonium, MD: York Press.

Vihman, M. M. (1993a). Variable paths to early word production. *Journal of Phonetics*, 21, 61-82.

Vihman, M. M. (2010). Phonological templates in early words: A cross-linguistic study. In C. Fougeron, B. Kühnert, M. D'Imperio & N. Valée (eds.) *Laboratory Phonology 10* (pp. 261-284). New York, NY: Mouton de Gruyter.

Vihman, M. M. (2014). *Phonological development: The first two years*. (2nd ed.). Oxford: Blackwell.

Vihman, M. M. (2015). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition* [FirstView Article], 1-20.

Vihman, M. M. & Croft, W. (2007). Phonological development: toward a 'radical' template phonology. *Linguistics*, 45, 683-725.

Vihman, M., DePaolis, R. A. & Keren-Portnoy, T. (2009). A dynamic systems approach to babbling and words. In E. Bavin (ed.), *Handbook of Child Language* (pp. 163-182). Cambridge, UK: Cambridge University Press.

Vihman, M. M. & Keren-Portnoy, T. (eds.). (2013). *The emergence of phonology: Whole-word approaches, cross-linguistic evidence*. Cambridge, UK: Cambridge University Press.

Vihman, M. M., Keren-Portnoy, T., Whitaker, C., Bidgood, A., & McGillion, M. (2013). Late talking toddlers: Relating early phonological development to later language advance. *York Papers in Linguistics*, 2, 47-68.

Vihman, M. M. & McCune, L. (2001). Early phonetic and lexical development: a productivity approach. *Journal of Speech, Language and Hearing Research*, 44, 670-684.

Vihman, M. M., Nakai, S., DePaolis, R. A. & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50, 336-353.

Vihman, M. M. & Velleman, S. L. (1989). Phonological reorganization: A case study. *Language and Speech*, 32, 149-170.

Vihman, M. M. & Vihman, V-A. (2011). From first words to segments: A case study in phonological development. In I. Arnon & E. V. Clark (eds.), *Experience, Variation, and Generalization: Learning a first language* (pp. 109-133). TiLar 7. Amsterdam: John Benjamins.

Vogt, S. & Kauschke, C. (2015, March). *Iconic gesture as a resource for word learning in typically developing children and children with specific language impairment.* Paper presented at the 37th Deutsche Gesellschaft für Sprachwissenschaft conference, Leipzig, Germany.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Waterson, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics*, 7, 179-211.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3, 1-30.

Werker, J. F., Gilbert, J. H., Humphrey, K. & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349-355.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103, 147-162.

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.

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

Wescott, R. W. (1971). Linguistic iconism. *Language*, 47, 416-428.

Westermann, G. & Miranda, E. R. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, *89*, 393-400.

Whalen, D. H., Levitt, A. G. & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French-and English-learning infants. *Journal of Child Language*, 18, 501-516.

White L., Floccia C., Goslin J. & Butler J. (2014) Utterance-final lengthening is predictive of infants' discrimination of English accents. *Language Learning*, 64,
27- 44.

Whitney, W. D. (1867). *Language and the study of language: Twelve lectures on the principles of linguistic science*. New York, NY: Scribner's.

Winter, B. (2013). *Linear models and linear mixed-effects models in R with linguistic applications*. arXiv:1308.5499. [http://arxiv.org/pdf/1308.5499.pdf]. Last accessed 16/07/2015.

Woodward, A. & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, 70, 65-77.

Yont, K. M., Snow, C. E. & Vernon-Feagans, L. (2003). The role of context in mother–child interactions: An analysis of communicative intents expressed during toy play and book reading with 12-month-olds. *Journal of Pragmatics*, 35, 435-454.

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S. &

Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, 124, 342-349.