Laterality and Babble:

Does asymmetry in lip opening during babble indicate increasing left hemisphere dominance as babies gain articulatory experience?

Florence Aithne Rose Oxley

PhD

University of York

Language and Linguistic Science

October 2022

Abstract

Speech and language are supported by task-dependent neural networks that are predominantly lateralised to the left hemisphere of the brain, whilst emotion is supported by predominantly right hemispheric networks. This is reflected in the asymmetry of lip openings during speech and facial expressions in adults. One cross-sectional orofacial asymmetry study found an analogous distinction between 5-12-month-old babies' lip openings during reduplicated babble and during positively valenced emotional facial expressions and this has been interpreted as evidence to support the hypothesis that babble is fundamentally linguistic in nature (Holowka & Petitto, 2002). However, a similar distinction is also observed in orofacial behaviours in some non-human primates. Differential hemispheric specialisation for emotional and vocal communicative functions may then be an ancient trait, long predating human language. Additionally, characterising babble as babies' immature attempts to do language marginalises the critical role of endogenously motivated vocal exploration and may assume a degree of goal-directedness in infant behaviour around the time of babble emergence for which we have little other supporting evidence.

This thesis explores laterality in eight 5-12-month-old's babble, positive facial expressions, and other vocalisations longitudinally. Singleton and variegated babble are captured as well as reduplicated babble, and an alternative method for analysing orofacial asymmetry – hemimouth measurement – is used. Overall, Holowka and Petitto's between-category distinction was replicated. However, babble was found to show right laterality at emergence and become left lateralised gradually over developmental time. Some interactional effect of utterance complexity was also observed. Bisyllabic babbles showed significant leftward shift over developmental time, whilst monosyllabic and polysyllabic babbles did not. Furthermore, hemimouth measurement revealed a degree of real-time variability in the laterality of babble not previously observed. An alternative theory of the underlying nature of babble – the Old Parts, New Machine hypothesis – is proposed.

Table of contents

List of tables

List of figures

Acknowledgements

This thesis represents the culmination – but, I hope, not the end – of a long and formative journey and thanks are due to the people who have helped me get here.

First and foremost, I extend my sincerest thanks to my supervisors, Prof. Tamar Keren-Portnoy and Dr. Eytan Zweig, without whose unwavering support, encouragement, insight, patience, and friendship, this thesis and the project behind it simply could not have happened. I thank them for every piece of practical and theoretical advice and guidance from wearing good socks for home visits, to working with mixed effects models, to developing my writing. I also thank Prof. Marilyn Vihman for her instruction and guidance during my Master's degree, which formed the foundation to this project. I thank them all for introducing me to the theories and ideas that have prompted a 'reorganisation' in the way I think about development and evolution. My time as a postgraduate student has been quite transformative and I am a much better researcher and writer for their tuition.

The incredible generosity and commitment of the York families involved in this study must also be acknowledged. They welcomed me into their homes with enthusiasm and plenty of cups of tea. I was quite sorry when the home visits came to an end. Acknowledgement is also due to the research assistants who performed reliability checks on each part of my methodology. I thank them for their diligence and their interest in the work. Thanks also to the Maths Skills Centre, Shayne Sloggett, and Vincent Hughes for sharing their expertise in statistical modelling and R Studio.

I also thank my employers and colleagues at the University of Leeds, chiefly Prof. Alice Deignan, Dr. Duygu Candarli, Dr. Dogus Oksuz, Dr. Robbie Love, and Fiona Scarth. I thank them for their guidance and for a wonderfully unexpected string of challenges and opportunities for professional and academic development. I have learned a lot from them too.

I also owe my thanks to my family and friends for their support and encouragement. Thanks to my mother, Karen Oxley for her insights as a paediatric Speech and Language therapist and for encouraging me to maintain some semblance of a work-life balance. Thanks to my father, Stewart Ellinson, for sharing his interest in sociology and psychology with me and for

consistently and insistently believing in my ability to do things even when I am not terribly convinced myself. Thanks to my cousins and my siblings. Watching them grow and change has, I am sure, in no small way sparked my interest in the sounds babies make and why they make them. Thanks to my sister Beatrice Ellinson, who asks often about my work despite claiming not to understand it, and who called me [dat] (‘that’) for a good year or so before learning my name. Now I know why. Thanks to my brother Dr. Barnaby Roper for 28 (almost 29) years of friendship, for chats about PhD life, and for at least trying with [fɒfəʊ] (‘Foffo’). Thanks also to Charlotte Wynn for her friendship and encouragement, and for talking statistics with me.

I also thank Zoe Pickburn for sharing her insights and personal reflections on pregnancy, motherhood, and pre- and post-natal development, and to Violet Pickburn-Tordoff for her part in that and some insights of her own. I look forward to working with them both on the next project. Thanks to the parents and little ones of Regia Anglorum who have shared many interesting and moving tales and adventures with me. Thanks to Fernando Pfaltz for his friendship and his support in accommodating my deadlines. Thanks to the other friends I have made along the way for sharing parts of the journey with me – Jared Stoughton, Sarah Wilson, David Shaw, Alex Foote, Daniel Keers, and Alex Hudson.

Lastly, I would like to thank Jordan Hayward, though I am not sure where to start. They have kept me grounded and we have spent many happy hours talking about evolution. They have seen most of the journey in fragments from behind the scenes but have never lost interest or patience and have always been somewhere in the wings with the right words to say in case I faltered. My resident archaeologist and my invaluable prompt.

Post-script: Very lastly, I thank Ulla the cat, whose company and keen interest in typing have been a little bit of a distraction and a lot of a joy while I have been making minor corrections to this thesis.

Author’s Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as references.

Some of the material in Chapters 1, 2, and 5 was included in brief in poster presentations at the International Child Phonology Conference (Crete, June 2018), the Lancaster Conference on Infant & Early Child Development (Lancaster, September 2018), and Building Linguistic Systems (York, June 2022). Some of this material was also included in a talk presented at the Humanities Research Centre Doctoral Fellowship Finals (York, June 2019).

Introduction

This introductory chapter will open by introducing babble and discussing how babble is defined in the context of phonological development. The theoretical perspective taken in this thesis – Dynamic Systems Theory – will be introduced, and the motivation for taking this perspective will be explained. The remainder of the chapter will then outline the structure and intentions of this thesis.

The language of babble: defining terms

Colloquially, the term *babble* has been used since around the thirteenth century to refer to fast, “foolish”, or emotional speech that is meaningless or indecipherable ("babble, n.," n.d.). In the context of phonological development, however, *babble* denotes a more specific behaviour – namely, babies’ production of meaningless adult-like syllables, typically emerging during the latter half of the first year of life (Oller, 1980; Stark, 1980; Vihman, 2014). Babble is not speech, but it constitutes the most ‘speechlike’ behaviour that babies do during the first year and this makes it a particularly salient phonological milestone to their caregivers and those around them (Oller, 1980; Oller et al., 2001; Fagan, 2009; Vihman, 2014). Phonological development and first language acquisition theory and research have, largely, used consistent criteria for identifying babble, with some discrepancies. This thesis assumes Oller and colleague’s criteria (Oller et al., 1976; Oller, 1980; Oller & Eilers, 1988; Oller, 2000). Under these criteria, babble is described as modally voiced alternation/s between at least one supraglottal closure (a consonant or consonant-like portion) and at least one vocalic opening (a vowel or vowel-like portion), produced without meaning or reference. These syllables are produced voluntarily and with mature articulatory control, resonance, and timing of the transition between the closure and the opening of the vocal tract (Oller, 1980; Stark, 1980; Oller & Eilers, 1988; Oller, 2000; Nip et al., 2009; Esling, 2012; Vihman, 2014). Oller’s criteria are widely accepted and are adopted here for consistency with the majority of the previous research in the field.

Whilst babble is commonly described as ‘speechlike’ or ‘languagelike’, babble is different from speech and language in some critical ways. Firstly, Oller (1980) has argued that it is important to reflect on what constitutes ‘speechiness’ and to clarify what is meant by ‘speechlike’. He suggests that ‘speechiness’ can be assessed using two frameworks – a *concrete* phonological framework and a *metaphonological* framework. Babies’ earliest vocalisations can be described only in terms of metaphonological features like pitch, phonation type (voice quality), resonance, timing, and amplitude. However, when babies begin to produce mature-sounding syllables, their vocalisations, like adult speech and language, can be described in terms of both these metaphonological and concrete features like stress, vowel height and rounding, or place and manner of articulation (p. 93). Since babble shares some of these metaphonological and concrete properties, it can be said to be somewhat speechlike. However, as we will see in Chapter 1, babble often shows some phonetic and phonotactic differences from the speech and language to which babies are exposed (DePaolis et al., 2011, 2013; Majorano et al., 2014). Secondly, and critically, babble is not used to encode reference or convey meaning (Oller, 1980; Vihman, 2014). Babble typically emerges around 6-8 months of age, some time prior to the capacities for processing symbolic reference and intentional communication (Oller et al., 1976; Oller, 1980; Iverson et al., 2007; Nip et al., 2009; Vihman, 2014). Whilst there is some degree of individual variation in babies’ age at the onset of babble, this age range is “robust” across both full-term and pre-term babies and across socioeconomic and linguistic backgrounds worldwide (Oller, 2000; Vihman, 2014, p. 30). Like other vocal-motor behaviours (e.g., cooing), babies sometimes produce babble during interactions with those around them (Meins, 1998; Vihman, 2014). Babble, like cooing, may therefore be involved in developing an understanding that vocalisation can be used for social interaction. However, like cooing, babble is not used as an arbitrary signal to intentionally encode semantic information. Furthermore, some babies have been observed to produce more babble *outside of* interaction, during episodes of solitary play (Vihman et al., 1985; Oller et al., 2019). This point will be revisited in more detail in Chapter 2.

A word on Dynamic Systems theory

Oller's (1980) criteria for babble are also preferred in this thesis for their compatibility with the Dynamic Systems theory perspective on development in biological organisms (Smith & Thelen, 1993; Thelen & Smith, 1994). Dynamic Systems theory offers a conceptualisation of behaviour and development in biological organisms that can be applied consistently across functional domains and levels (from the cellular level, to organs, to the body, and the environment) and across species (Smith & Thelen, 1993; Thelen & Smith, 1994; King & Shanker, 2003; Mareschal et al., 2007). Behaviour is theorised to become increasingly sophisticated in a time- and experience-dependent manner through cyclical interactions between the organism, itself, and its environment, which includes other organisms of the same or other species (Thelen, 2005; Mareschal et al., 2007). This comes about through cycles of global self-organisation, resulting in periods of more homogenous, stable behaviour that are occasionally disrupted and altered by the introduction of some lower-level *chaos* or instability (Thelen, 1989; Smith & Thelen, 1993; Thelen & Smith, 1994). Dynamic Systems theory draws on principles of Chaos Theory. Evolution and development in dynamic biological systems come about through changes which may seem superficially to be random and unpredictable, but which are in fact determined and constrained by the organism's dynamic history, and will go on to further constrain the future development of the organism (Smith & Thelen, 1993; Thelen & Smith, 1994; Mareschal et al., 2007).

Dynamic Systems theory paints a picture of babble and other newly emerging behaviours in infancy and childhood as *systems* formed from confluences of diverse contributing skills and behaviours or *subsystems* (Thelen & Smith, 1994). The terms *system* and *subsystem* are used to capture the coherent, regular, and repeatable co-ordination or synergy of movements of different parts of the body in behaviours like reaching, stepping, or babble (Thelen & Smith, 1994). The subsystems of any given system often operate in multiple generalised domains. With regard to babble, these include the motoric, vocal, rhythmic, social, sensory-perceptual (auditory, visual, and proprioceptive) and cognitive domains. Subsystems become associated or *entrained* when they are repeatedly co-activated across multiple instances (Thelen & Smith, 1994). Treating babble as a system accounts for the way that it manifests from the bottom up as a globally stable form of behaviour, which may vary across instances of production, and which may change non-linearly over developmental time when changes occur at the level of its subsystems (Thelen, 1989; Smith & Thelen, 1993; Friend, 2004; Gershkoff-Stowe & Thelen, 2004; Goldin-Meadow, 2004; Marcovitch & Lewkowicz, 2004).

Development in Dynamic Systems theory is, critically, not viewed as a teleological phenomenon (Thelen & Smith, 1994; Mareschal et al., 2007). Each form of a developing behaviour represents a developmentally complete form, emerging from pre-existing behaviour/s, rather than a discrete, stepwise move towards some pre-determined endpoint. Subsystem-level changes are prompted by destabilising variables or *control parameters*. These introduce some instability to the foundations of a stable behaviour. This provokes reorganisation of the overall system and this reorganisation, in turn, results in the emergence of a new stable system (Thelen, 1989; Thelen & Smith, 1994; Gershkoff-Stowe & Thelen, 2004). In development, subsystem-level changes may occur when babies acquire a new skill that may come to be incorporated into a pre-existing system as a new subsystem, or when they become more proficient at producing a behaviour that already represents a subsystem of a given system or systems. Babble and its precursors draw on multiple heterogeneous subsystems that develop asynchronously, so new endogenous control parameters emerge relatively frequently, giving rise to babies' rapid and successive attainment of phonological milestones during the first year of life. System reorganisations indicate an increase in the sophistication of a system. The emergence of different subtypes of babble, the shift from shorter to longer babble utterances, the establishment of babies' favoured consonants and idiosyncratic phonological systems, and of their first word form productions represent some particularly salient behavioural reorganisations that can be observed during the first year or so, as babies gain articulatory practice and new cognitive and motor skills like the capacities for articulatory planning and symbolic representation (Oller, 1980, 2000; Thelen & Smith, 1994; Vihman et al., 2009; Vihman, 2014).

Oller's (1980) criteria and his accompanying discussion account satisfactorily for the dynamicity and idiosyncratic variation observed in babble and wider vocal development, aligning well with a view of phonological development as a series of loosely chronological transitions between milestones, which may sometimes occur simultaneously or overlappingly, and which may appear superficially to show regression in sophistication at times (Thelen & Smith, 1994; Vihman, 20014, 2019). Alternative characterisations of phonological development have promoted a view of development in babble and word acquisition as a chain of discrete, stepping-stone-like stages, moving linearly from less to more sophisticated (e.g., Stark, 1980). Such characterisations are limited in that they do not offer such psychologically, biologically, or neurologically plausible (non-teleological) explanations for how and why babies move between 'stepping-stones'; do not account for apparently regressive behaviours; and – importantly – are not consistent with wider theories of development in biological organisms (Thelen & Smith, 1994; Mareschal et al., 2007). Parsimony being desirable in scientific theory (Smith & Thelen, 1993), the Dynamic Systems theory perspective is preferred here.

Why research babble?

As we will see in Chapter 1, babble seems to be a universal and necessary precursor to language acquisition. Understanding the nature of babble – where it comes from, why we do it, and how it shapes our subsequent development – can offer important insights into the ontogeny and phylogeny of language. Chapter 1 will describe how a confluence of vocal, rhythmic, and motor behaviours become entrained in babble, and will explore how babble may later be combined with other capacities when babies begin to acquire language, from the perspective of the hypothesis on babble taken in this thesis – the *Old Parts, New Machine* hypothesis. Chapter 1 will also explore babies' intrinsic motivation to babble – their sense of enjoyment in experimenting with articulation, cause and effect, and social interaction with those around them. Chapter 2 will discuss other competing hypotheses about the developmental and evolutionary origins of babble alongside some evidence from cross-species comparison and atypical phonological development. Chapter 3 will discuss the parts of the brain involved in linguistic and pre-linguistic vocal and emotional orofacial behaviours and their possible evolutionary origins, as observed through human and non-human research. Working to understand the processes and mechanisms involved in the emergence and production of babble can offer valuable insights into the extent to which babble and, by extension, speech, are governed by biological or culturally mediated, extrinsic or intrinsic factors, and, ultimately, how the human capacity for language evolved.

Typically and atypically developing babies from all cultural, socio-economic, and linguistic backgrounds produce at least some babble before acquiring language (Locke & Pearson, 1990; Bleile et al., 1993; Davis & MacNeilage, 1993, 1995; Alcock, 2006; Bohm et al., 2010; Vihman, 2014; Patten et al., 2014). Whilst certain differences can be noted between the babble of typically and atypically developing babies, the fact that babble represents a universal phase in phonological development suggests that babble plays a critical role in phonological development and that it may be at least partly the product of some intrinsic impulse (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; Elbers, 1982; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; Locke & Pearson, 1990; Bleile, 1993; Lenhoff et al., 1997; Locke, 2000; Masataka, 2001; Stoel-Gammon, 2001; Keren-Portnoy et al., 2005; Alcock, 2006; DePaolis et al., 2011; Whitworth & Bray, 2015; McGillion et al., 2017). The incidence of atypical babble development represents a further factor motivating babble research. Some babble research has already had clinical and diagnostic applications (e.g., Masataka, 2001; Stoel-Gammon, 2001; Alcock, 2006; Whitworth & Bray, 2015). Babble provides babies with the phonological foundations for communicating using spoken language (Vihman & McCune, 2001; Keren-Portnoy et al., 2005; McGillion et al., 2017), and research has shown that certain atypicalities in babble are correlated with diagnoses associated with delayed or disordered language development later in childhood, like Autism Spectrum Conditions, hearing impairment, or Developmental Verbal Dyspraxia, amongst others. Identifying features of atypical babble and their short- and long-term consequences for development could facilitate early identification of babies who may require additional support. Some developmental differences that are associated with atypical speech and language development, such as Down's Syndrome, Edwards' Syndrome or Patau's Syndrome, can be identified at or before birth (NHS, 2019). Working to understand how babble in these cases differs from typical babble can inform how we construct therapies and interventions for children with differences like these and can help us to understand how and when to implement these most effectively. It is widely observed that outcomes are significantly better where diagnoses and targeted support are available as early in life as possible. To learn about how typical and atypical babble differ, it may be helpful to build a comprehensive picture of the characteristics of typical babble as a starting point.

Much research has already been conducted on the phonological and articulatory-phonetic properties of typical and atypical babble and the other early vocal behaviours preceding and cascading from babble (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; Elbers, 1982; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; Locke & Pearson, 1990; Bleile, 1993; Lenhoff et al., 1997; Locke, 2000; Masataka, 2001; Stoel-Gammon, 2001; Keren-Portnoy et al., 2005; Alcock, 2006; DePaolis et al., 2011; Whitworth & Bray, 2015; McGillion et al., 2017). Some research has also examined typical and atypical socio-emotional and communicative development, and investigated the neural activity and development that underlies this (Nicholls et al., 2005; Raja Beharelle et al., 2010; Bishop, 2013; Corballis, 2013; Szaflarski et al., 2014; Riès et al., 2016; Lindell, 2016; 2020 Cantiani et al; 2019). Very little research, however, has investigated the neural activity that underlies babble specifically (Holowka & Petitto, 2002a). Since all behavioural development is underlain by adaptive pruning and strengthening – and formation – of networks of neurons and brain areas (Johnson, 2000; Rosselli et al., 2014; Williamson & Lyons, 2018), such research could be of great value in working to understand the ontogeny and phylogeny of babble and language. The present study advocates using hemimouth measurement as a non-invasive, rigorous, and practicable method for analysing the laterality of infant vocal behaviour and emotional facial expressions, that is sensitive to subtle changes in brain development over time (see Schuetze & Reid, 2005; Oxley et al., 2014). This method will be used, in this study, with typically developing infants. Future research could apply this method with babies from special populations to test for between-population differences in the laterality of babble.

The present study

The remainder of this thesis will present the research carried out in this project and a discussion of its findings. Hemimouth area measurement is used to examine the laterality of babies' babble, other non-babble vocalisations, and positively valenced emotional facial expressions (smiles and laughs). The present study uses a modified version of a method first developed by Hook-Costigan and Rogers (1998) and later adapted by Fernández-Carriba et al. (2002a, 2002b), Losin et al. (2008), Wallez and Vauclair (2011, 2012), and Wallez et al. (2012). The method is further adapted here for rigour and for suitability for use with human infant video data. The resulting methodology is presented in Chapter 4, along with a description of the participants recruited and the data generated in this study. In Chapter 5, the results of the statistical analyses of these data are presented. Mixed effects models are used, having the advantage that they are able to distinguish between the effects of both random and fixed predictors, can handle large quantities of data as were generated by this research, and can accommodate missing values in repeated measures – a typical phenomenon in naturalistic research with babies – without sacrificing statistical power. Chapter 6 will reflect on the findings of this study and will discuss their implications for future research and the contributions of this research to the fields of laterality and phonological development.

Chapter 1 Babble and Language

This chapter opens by discussing in more detail the criteria used for identifying babble in this and other studies. The chapter will then go on to discuss how babble emerges from pre-existing behaviours in a baby's repertoire, how it becomes more sophisticated with time and articulatory practice, and how babies move from producing babble to beginning to produce their first words. The following chapter will discuss the debate surrounding the underlying nature of babble.

"[T]he cognitive machinery that makes us human can be viewed as a new machine constructed out of old parts"

(Bates, 1999, p. 10)

Introduction

This thesis puts forward the hypothesis that babble emerges endogenously as a confluence of some of the pre-existing behaviours and capacities in a baby's repertoire. Under this hypothesis, babble and its emergence are shaped and constrained by what babies have done before and will itself shape and constrain their phonological development in the future. This hypothesis draws on research from diverse but interrelated fields including phonological development, developmental psychology, and comparative and developmental psycholinguistics and neurolinguistics and aims to be consistent with well-established principles of development and evolution in biological organisms. To maintain the distinction between this and other hypotheses regarding the underlying nature of babble – namely the *Motoric hypothesis* (see Lenneberg, 1967; Jakobson, 1968; Davis & MacNeilage, 1994) and *Linguistic hypothesis* (e.g., see Petitto et al., 2004) – this hypothesis is termed here the *Old Parts, New Machine hypothesis* following Bates (1999, p.10), whose words encapsulate the way in which nature borrows from itself in 'inventing' new phenomena. Whilst Bates wrote these words with reference to language evolution, it is argued here that the same principle holds for the behavioural phases in phonological development that precede it. This chapter will set out how phonological development progresses during the first year or so of life from the perspective of this hypothesis.

Defining babble: criteria

As discussed in the Introduction to this thesis, babble is defined here as consonant-vowel syllables with adult-like temporal organisation and without meaning or reference, emerging around 6th-8th month of life (Oller, 1980; Vihman, 2014). These may be produced singly (singleton babble), repeatedly (reduplicated babble), or as a string of syllables with consonants of differing place or manner of articulation (Oller, 1980). As was also mentioned in the introductory chapter, certain discrepancies can be noted between Oller's criteria and those used in a minority of cases in the research literature. These discrepancies are important to note in cases where differing definitions of babble may influence methodological decisions or may give rise to apparent inconsistencies in the findings of related research. Oller and Eilers (1988) note that the term *babble* is sometimes extended to include all non-vegetative, voluntarily produced infant vocalisations, regardless of structure and phonetic content. Elbers (1982) distinguished between *single*, *repeated*, *concatenated*, and *mixed* babble. Elbers' repeated babble corresponds to Oller's reduplicated babble, and Elbers' concatenated and mixed babble represent subtly different subtypes of variegated babble. Oller's singleton babble and Elbers' single babble differ more significantly: while Oller's criteria state that babble must contain a consonant, Elbers' single babbles included syllabic consonants produced without a vowel and isolated vowels produced without a consonant. Overextending the term in this way can obscure the significant shift in babies' phonological behaviour when they develop the neuromuscular control to produce consonant-vowel transitions with adult-like timing. Petitto and colleagues have favoured a definition of babble which prioritises reduplicated babble but excludes singleton and variegated babble (Petitto & Marentette, 1991; Holowka & Petitto, 2002; Petitto et al., 2004). The reasoning for this exclusion is not explained. While taking reduplication to be a defining feature of babble may facilitate faster collection of more homogenous samples of babble data, using this criterion risks representing babble as static or monolithic. Excluding singleton and variegated babble from babble research means that research findings may not fully represent the dynamicity and variability that is characteristic of babble, both within and across babies, and over real and developmental time. Oller's criteria are preferred here as they capture this dynamicity and variability but maintain the important distinctions between babble and the less mature vocalisations that precede it.

The remainder of this chapter will now explore how babble emerges through the entrainment of its underlying subsystems, how it develops qualitatively over time, and how it contributes to later language acquisition.

Phases within babble

Taking this view of babble as a dynamic system, this next section will outline the behaviours that come to be bound together to form babble. Oller (1980) characterises these behaviours as eponymous "stages", though his description of these stages and babies' transitions between them is perhaps better captured by the term *phases*. Kent (1984) has argued that a process model might capture continuities in phonological development more faithfully than a stage-based model. However, Oller (1980) offered the caveat that the stages that he described were neither necessarily suddenly emerging nor discrete – rather, stages may overlap, may emerge and/or decline gradually, and may appear superficially to show regression in sophistication. Oller also notes that these stages and the order in which they emerge are not necessarily exhaustive nor universal. Several authors have described similar stage models in phonological development and have echoed Oller's caveats (Stark, 1980; Studdert-Kennedy, 1990; Stoel-Gammon, 1992). Stoel-Gammon (1992) has added a further caveat: that the new behavioural form from which each stage takes its name may not always be the behaviour most frequently produced during that stage. Studdert-Kennedy (1990) outlines four stages spanning 0-24 months. However, while the age ranges at emergence of some behaviours described correspond roughly to those put forward by Oller (1980), Studdert-Kennedy's stages group all vocalisations preceding babble into a single category, and group variegated babble into a category with first word acquisition. The distinctions between precursor behaviours to babble, and between variegated babble and speech are of interest to this thesis. As such, Studdert-Kennedy’s stage model is not discussed in further detail here. While Stark's (1980) and Stoel-Gammon's (1992) stages differ only slightly from Oller’s in terms of ages ranges and terminology, these authors describe largely the same behaviours, emerging over the same timescale. For brevity and consistency, therefore, this thesis will focus primarily on Oller (1980) stage model from hereon, with brief reference to other models and research where relevant.

Oller (1980, p. 95) states that the transitions between stages “represent[s] the incorporation of some new non-reflexive vocalisation characteristic(s) into an infant’s controlled repertoire” (see also Stark, 1980). In Dynamic Systems terms, transition between phases happens when a baby acquires, or attains greater proficiency in, a metaphonological or concrete phonological ability, prompting a reorganisation of pre-existing stable behaviours that use part/s or all of the same articulatory apparatus or capacities. So, rather than discrete, universal steps, the stages described below represent general tendencies amongst babies for these behaviours to emerge around certain ages (assuming typical development[1]), and to be produced with sufficient frequency and consistency to indicate voluntarily controlled production. Aligning closely with the work of other researchers in phonological development and in general development in biological organisms, Oller’s (1980) stage model is taken here to offer a satisfactory synopsis of key phonological milestones attained during the first year of life, while accounting for the substantial variability typically seen in infant data. However, as argued here, the rigidity and discreteness implied by the term *stage* does not fully reflect the dynamicity of phonological development and obscures the flexibility of Oller’s stages as he describes them. The term *phase* is, therefore, used in this thesis rather than *stage* to better capture the continuous, transitional nature of development and the sometimes-fuzzy boundaries between the onset and decline of each different behaviour.

---

[1] See Chapter 2 for a discussion of some cases of babble and atypical development

The Phonation Phase (0-1 month)

During the first two months of life, babies produce very few sounds that can be considered 'speechlike' (Oller, 1980; Vihman, 2014). Newborns' vocalisations are restricted to reflexive sounds like discomfort or distress cries and vegetative sounds like eructation ('burping') and swallowing (Oller, 1980; Vihman, 2014). By 6-8 weeks, babies have begun to vocalise when in a contented affective state (Stark, 1980). During this time, through repeated experience with crying and 'haphazard' reflexive phonation (e.g., fussing), babies acquire the ability to initiate vocal fold vibration intentionally (Oller, 1980). Later, after around 8-12 weeks of life, babies may begin to extend intentional phonation to voluntarily produced "comfort sounds" like cooing and laughing (Stark, 1980; Vihman, 2014, p.82). These voluntary sounds are termed *Quasi Resonant Nucle*i (*QRNs*) by Oller (1980; Oller & Eilers, 1988). Quasi Resonant Nuclei are low-amplitude, low-resonance (<12000Hz), modally-voiced sounds, which may be vowel-like elements, syllabic consonants (including stops, clicks, fricated sounds, and trills), or "small throaty noises", and which are not produced with any systematic opening or closing of the vocal tract (Oller, 1980, p. 85; Stark, 1980). During this time, some babies do also produce a small proportion of more open vowel-like sounds with resonance >12000Hz – or, in Oller's terms, *Fully Resonant Nuclei* (*FRNs*) – but these do not yet occur frequently enough to suggest that they are voluntarily produced. Recall that frequent production is a widely used criterion for assuming that a behaviour is voluntarily produced and well-established in a baby's repertoire. Quasi Resonant Nuclei are produced with an "at-rest […] position of the supraglottal vocal tract" whereas during Fully Resonant Nuclei, the vocal tract is more open (Nathani Iyer & Oller, 2008, p. 7).

Phonation and supraglottal modulation of the airstream are secondary functions of the vocal tract – its primary function being to protect the lungs from obstruction and to perform feeding and swallowing (Koopmans-Van Beinum et al., 2001). In the early weeks of life, aside from crying, babies from across language backgrounds have been found to produce more epiglottal and pharyngeal sounds than other sounds (Esling, 2012 c.f. Oller's (1980, p. 85) "small throaty noises"). These back closures or strictures are close in articulatory form to closing the pharynx to protect the airway – an innate, reflexive movement triggered by the brain stem (Esling, 2012). Babies' voluntary production of these back sounds demonstrates their acquisition of metaphonological skills from reflexive vocalisations, and the reorganisation of these skills into a new, globally stable behavioural form – a voluntarily produced Quasi Resonant Nucleus. The sensory information for creating phonation and vocal tract closure, initially experienced by babies as the consequences of reflexive movements and their immature neuromuscular control over these (Vihman, 2014), becomes associated and embedded in the productive repertoire. Esling (2012, p. 134) also suggests that producing more "'open'" sounds could present some "risk". He does not specify what manner of risk, but he may be referring to accidental inhalation of saliva. Once established in repertoire, the movements for creating a sound become available for voluntary production, given sufficient motivation. This motivation may come about through babies' exploratory play: babies may enjoy hearing a certain self-produced sound or may enjoy the haptic experience of producing it. Whilst not yet speechlike, these low-resonance pharyngeal vocalisations provide babies with a catalogue of experiences of vocal fold vibration, and of voluntarily modulating the airstream from their lungs, resulting in phonation; two critical pre-requisites for later speech and language (Oller, 1980). It is suggested here that, while these Quasi Resonant Nuclei represent a developmentally complete system in themselves, through producing these sounds, babies' neuromuscular articulatory control strengthens, making them capable of producing movements more precisely and consistently. This enhanced precision and consistency eventually introduces enough instability into the subsystems of Quasi Resonant Nucleus production to trigger the further reorganisation, resulting in Fully Resonant Nucleus production, typically between 4-6 months of age.

The GOO phase and Turn-taking (2-3 months)

Around 2-3 months of age, infants begin to produce a high proportion of *GOO* sounds (Oller, 1980; Fagan & Iverson, 2007). GOO sounds are repetitive articulations comprising a somewhat imprecise closure or stricture between the back of the tongue (the dorsum) and soft palate (the velum), and a vowel-like *Quasi Resonant Nuclei* (Oller et al., 1976; Oller & Eilers, 1988). The frequency with which babies are observed to produce these sounds may be influenced by gravity and posture. At 2-3 months of age, many babies spend a good deal of time lying supine (Kent, 1984; Esling, 2012). When lying supine, the tongue rests closer to the velum than when sitting upright, meaning that closure between these articulators requires less movement and may even occur incidentally at times (Esling, 2012). Another explanation for the prevalence of velar articulations in this phase many be found in physiology. At birth, the infant head and vocal tract are relatively small, the neck is short, and the larynx is positioned high in the pharynx, while the tongue is proportionally large and is situated mostly in the mouth, proximal to the epiglottis and velum (Kent, 1984; Esling, 2012; Lieberman, 2017). In fact, the infant vocal tract is closer in morphology to that of non-human primates than to that of human adults or even older children (Oller, 1980; Stark, 1980; Kent, 1984). Only around 50% of the tongue occupies the mouth in adults, while the remaining 50% lies vertically in the throat (Lieberman, 2017). These physiological differences limit the range of articulatory gestures and phonation types available to babies and this contributes to the relatively high proportion of 'back' sounds observed in the Phonation and 'GOO' phases. Around 2-3 months of age, the physiology of the respiratory and vocal tracts grows and changes, and babies develop stronger neuromuscular control (Lieberman, 2017). During the GOO phase, babies are able to sustain voicing for longer and produce articulatory gestures more consistently (Bloom, 1998). With a proportionally smaller tongue size and a greater distance between the dorsum, the velum, and the epiglottis, babies are now able to produce a wider range of back sounds (Bloom, 1998). Esling (2012) also notes that glottal and, by extension, velar sounds require more sophisticated neuromuscular control than epiglottal and pharyngeal sounds.

It is around this time that babies also begin to show a cognitive attentional shift and may be observed to engage in 'turn-taking' with their caregivers (Bloom, 1998; Vihman, 2014). Babies now attend closely to the movements of caregivers' eyes and faces and the sounds that they produce, and respond by moving their own mouths, with or without phonation (Bloom, 1975; Bloom, 1977; Bloom, 1998).

Oller (1980, p. 96) notes that Zlatin (1974) has suggested that GOO sounds may show "primitive syllabification". However, Oller disputes this claim, remarking that any transitions between open and closed portions of sounds produced during this phase do not yet show the advanced neuromuscular control, nor the temporal organisation that characterises mature or even immature syllable production. That is not to say, however, that the GOO phase does not represent a significant milestone in phonological development. During this phase, physiological and neuromuscular changes introduce enough instability into the subsystems underlying Quasi Resonant Nucleus production to prompt the reorganisation that results in GOO production. Modal phonation, originating in vegetative sounds and becoming voluntary during the phonation phase, can now become coupled with a wider possible range of places of articulation at the back of the oral tract (Oller, 1980; Kent, 1984; Esling, 2012; Vihman, 2014). It is suggested here that cognitive attentional advances may introduce further instability in the form of a new motivational drive – a drive to attend to caregivers' behaviour (Mundy et al., 2000; Vihman 2014) – which may also influence the reorganisation and reorientation of GOO sounds. Originally a purely endogenous vocal motoric behaviour, babies' vocalisations can now sometimes be elicited by social interaction (Bloom, 1975; Bloom, 1977).

The Expansion Phase: marginal babble and motor planning (4-6 months)

During this next phase, babies' vocal repertoires expand exponentially as they engage in *vocal play* (Stark, 1980), experimenting with different configurations of metaphonological features (pitch, amplitude, resonance, and timing) (Oller, 1980). Between age 4-6 months, babies may begin to squeal, growl,[2] yell, 'blow raspberries',[3] produce alternating ingressive-egressive sequences, and – perhaps most significantly in the context of phonological development and language acquisition – may begin to produce *marginal babble*. Marginal babble comprises alternations between at least one Fully Resonant Nucleus and at least one closure or stricture in the vocal tract (Oller, 1980). Oller uses the label *marginal* as these vocalisations are not fully-formed babble; they do not have the adult-like temporal organisation that characterises c*anonical babble*. However, the resonance of the vocalic portion shows more sophisticated neuromuscular control than the vocalic sounds produced during the GOO phase. Marginal babble emerges through reorganisation of the pre-existing system (the GOO sounds), triggered when babies become more proficient at producing and maintaining resonance.

It is also around this time that babies begin to show signs of an emerging ability to prepare motor movements before producing them (Vihman, 2014). This capacity for motor planning is domain-general, with consequences for gross and fine manual motor skills as well as oral

---

[2] Creaky voice

[3] Bilabial or linguolabial trills and vibrants (Oller & Eilers, 1988)

motor skills. This capacity likely represents a new subsystem, which also contributes to the reorganisation of the subsystems supporting GOO sounds and, thus, to the emergence of Marginal babble as a new system. The ability to plan articulatory sequences can, for example, give babies the ability to form articulatory gestures before beginning to vocalise, or allow them to plan to repeat more sophisticated vocalisations after recognising or remembering the link between an oral gesture they have previously made and the sound that was created.

Marginal babble is produced frequently enough to indicate that it is produced voluntarily, though it is produced relatively infrequently compared with other precursor behaviours like cooing, vocalic shouting, or experimenting with voice quality and supraglottal occlusions (e.g., squealing, growling, blowing raspberries) (Oller, 1980). Marginal babble may therefore represent a less stable, more transitional system which emerges and destabilises relatively rapidly. A combination of two control parameters may account for its rapid emergence and decline. Interestingly, Vihman (2014, p. 26) writes of this period that:

"[Nevertheless,] reaching and grasping are manifested in clumsy attempts at best at this age; they will be performed smoothly only from about 6 months on (Thelen, Corbetta & Spencer, 1996). According to Bruner (1973), this is the result of the initial action components being reorganised into a higher-order routine after a considerable period of practice."

Here, there is a clear parallel between manual motoric development and vocal motoric development: this period may represent a 'practicing' period wherein movement in multiple domains is initially immature and imprecise but becomes more sophisticated and controlled with accumulated productive experience. Articulatory practice, then, may represent one control parameter that triggers the shift from marginal to canonical babble. The term *practice* is not used here in the sense of intentional 'work' done by a baby to attain a specified goal. Rather it simply denotes repeated or habitual action, resulting in stronger neuromuscular control, which will contribute to and constrain the way in which the baby acquires and consolidates new skills in the future.

Soon after this time, the capacity for rhythmic movement begins to influence how babies vocalise (Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001). Being (typically) a relatively poorly established system by 6 months of age, marginal babble may be particularly susceptible to interference from rhythmic ability. The ability to move parts of the body (e.g., the limbs, extremities, head, or torso) by contrast, is much more well-established, having been in repertoire from around the first month of life (Thelen, 1979, 1981; Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001). Marginal babble involves articulatory planning and alternations between consonant- and vowel-like elements with adult-like resonance. When the capacity for marginal babble is made use of at the same time as the capacity for rhythmic movement, rhythmic ability may trigger a reorganisation of the timing of the transition between the articulatory gestures for these consonant- and vowel-like elements. How this may happen will be discussed later in this chapter.

The Canonical Phase: canonical babble (7-12 months)

Babies' arrival at the canonical phase, around 6-8 months of age, is marked by the emergence of voluntarily produced adult-like syllables (Oller, 1980; Oller & Eilers, 1988; Oller, 2000; Oller et al., 2001; Vihman, 2014). These syllables are formed from consonant-like supraglottal closures paired with oppositional vowel-like elements (Fully Resonant Nuclei), that are produced with adult-like resonance and neuromuscular control, and with consonant-vowel transitions that have "a timing relationship that conforms to mature natural language restrictions" (Oller, 1980, p. 98). Critically, canonical babble can be described in *concrete* phonological terms as well as metaphonological terms, marking a significant qualitative change in the babies' phonological behaviour; it is now more "speechy" (p. 98). Following the emergence of canonical babble, there are some further reorganisations in the subsystems supporting this new system. Canonical babble may be produced as single syllables (*singleton babble* e.g., [ba]), or as a string of syllables, containing either the same syllable repeated several times (*reduplicated babble* e.g., [nanana]) or a sequence of syllables containing different consonants (*variegated babble* e.g., [dagaja]). Singleton babble usually emerges earliest, while reduplicated babble typically emerges later (Oller, 1980). Variegated babble typically emerges latest – around 11-12 months according to Oller (1980) or between 10-14 months according to Stark (1980). These reorganisations reflect babies’ increasing neuromuscular control and proficiency with articulatory planning. Producing singleton babble requires babies to plan to articulate a single closure and opening in the vocal tract. Reduplicated babble requires slightly longer-term planning, and the co-ordination of repeated sequences requires some – likely implicit – awareness of order. Reduplication also requires babies to be able to move their articulators from closure to opening, and then back to closure whilst continuously phonating. Variegated babble requires still more sophisticated planning and sufficient articulatory control to move from an initial closure to an opening, and then to another closure in a different part of the vocal tract. Around 11-12 months, infants may also begin to experiment with stress patterns, producing utterances formed of a combination of canonical babble and imprecisely articulated syllables reminiscent marginal babble. This, Oller (1980, p. 99) terms “gibberish”.

With the emergence of babble, babies have the beginnings of the phonological resources that will later become involved in speech and language acquisition (Oller et al., 2001). Indeed, a number of researchers have noted that there is some considerable continuity and overlap in the phonetic structure of canonical babble and “nursery terms” like ‘mama’, ‘dada’, or ‘baba’ (for ‘bottle’) (Elbers & Ton, 1985; Oller & Eilers, 1988, p. 442; Vihman, 1993; Keren-Portnoy et al., 2005; Keren-Portnoy et al., 2010; DePaolis et al., 2011; DePaolis et al., 2013; Vihman et al., 2014). Relatedly, Mitchell and Kent (1990) have observed that the emergence of babies’ first words tends to coincide with a peak in the frequency of their reduplicated babble. So, then, canonical babble is often considered to resemble adult speech and language in ways that the behaviours that come before it do not. In this way, canonical babble is a ‘speechlike’ behaviour, composed of ‘unspeechy’ behaviours and capacities. Babble, like language and cognition, may then also represent a “new machine” made out of “old parts” (Bates et al., 1979; Bates, 1999, p. 10; Bates, 2004, p. 250) The following section will discuss these behaviours and capacities, and the roles that they play in further depth.

Pre-requisites for canonical babble

Koopmans-Van Beinum et al. (2001, p. 69) state that there are three necessary pre-conditions for the emergence of babble: 1), a “biologically governed” capacity for repetitive rhythmic movement; 2) sufficient neuromuscular control over the larynx and supraglottal architecture to produce voiced articulatory gestures intentionally; and 3) good quality auditory perceptual feedback. Under a Dynamic Systems Theory view, these pre-requisites become coupled together through repeated co-activation over many instances.

Rhythm

While rhythmicity may be more salient in gross motor movements like kicking or arm-flapping, it may act as a kind of behavioural *attractor* for other types of motor activity (Waddington, 1966, 1977, cited by Thelen & Smith, 1994). After rhythmic ability emerges, rhythmic movements gradually come to make up a high proportion of babies’ motor activity (Thelen, 1979, 1981). The frequent production of different rhythmic behaviours suggests that rhythmicity becomes a robust but flexible subsystem that is involved in a number of different systems. Thelen (1981) reports on no less than 35 different types of rhythmic behaviour emerging and peaking in frequency at different times between 4-52 weeks of age, and Thelen (1979) reports on 47. This stability and flexibility may allow rhythmicity to be extended to other motor behaviours, like vocalisation. Marginal babble, being less stable, may be especially susceptible to the influence of rhythmicity as a new subsystem. The emergent capacity for producing rhythmic behaviours or *stereotypies* (Thelen, 1979, 1981) may, then, ‘attract’ marginal babble and furnish it with new, more regular and adult-like rhythmicity, resulting in a reorganisation into a new system – canonical babble. By contrast, canonical babble typically comes to be a highly stable system, occupying many months and requiring some significant interference from later-emerging cognitive capacities to prompt the transition into word acquisition.

Rhythm research with babies has provided some valuable insights into how babble emerges, supporting the idea that babble originates endogenously, through exploratory play, rather than as a response to social interaction or to perceiving adult speech and language (Esling, 2012). Indeed, it has been found that babies babble more often whilst producing some other rhythmic movements or exploring their environment, than in the context of social interaction (Stark, 1980). Interestingly, non-human research has also found that the infants and juveniles of some bat, primate, and songbird species also produce more immature vocalisations outside of social contexts (Elowson et al., 1998; Knörnschild et al., 2006). These immature non-human vocalisations are sometimes also termed *babble* and represent a time when infants of these species are exploring the possible sounds that their articulatory apparatus can produce (Omedes, 1985; Elowson et al., 1998; Knörnschild et al., 2006; Knörnschild, 2014; Snowdon, 2018; Oller et al., 2019; Fernandez et al., 2021; Ter Haar et al., 2021). This research and its relevance to this thesis will be discussed in more depth in the final chapters of this thesis, though the term *vocal learning* will be preferred over *babble* to maintain the distinction between the specific behaviour investigated in this thesis and the vocal behaviours of non-human animals.

Articulatory control

How vocal motoric skills contribute to the emergence of babble is perhaps more transparent (Kent, 1984). Without having gradually accumulated experience of phonation and articulation during the first 6 or so months, which results in more sophisticated neuromuscular control, babies would not be equipped to begin to produce babble. In the early weeks of life, babies come to master voluntary phonation and often produce consonant-like laryngeal or velar strictures (Oller, 1980; Esling, 2012; Vihman, 2014). Later, babies have been observed to use these back sounds as a “prime” to “jump start” consonant-like articulations at the front and, later still, in the centre of the vocal tract (Elbers, 1982; Kent, 1984; Esling, 2012, p. 141). Esling (2012) further notes that babies whose ambient language contains laryngeals often continue to produce laryngeal sounds after oral consonants emerge, while in babies whose ambient language does not use laryngeals, these sounds tend to decline. For these babies, very early vocal experience may come to have a direct bearing not only on their ability to produce babble, but also on their language acquisition. It should be noted, though, that the ubiquity of laryngeal vocalisations in the earliest months in babies across language backgrounds suggests that this and other phonological behaviours before word acquisition are endogenously motivated.

As discussed previously, the prevalence of laryngeal and velar strictures may be explained by babies’ physiology and their body posture. How, then, might we account for the emergence, and later prevalence, of front and centre consonants in babble (Elbers, 1982; Esling, 2012)? Fagan and Iverson’s (2007) study of mouthing behaviour in 6-9-month-olds illustrates how babies’ experiences of mouthing their own hands, feet, and other objects contributes to their phonological development. Mouthing behaviour denotes babies’ exploration of objects in their environment, like toys or clothes, or parts of their own body, by contacting the object or body part with their mouth (Fagan & Iverson, 2007; Bukowska et al., 2010). The mouth hosts a particularly large number of nerve endings, meaning that babies gather especially rich sensory feedback through mouthing (Bukowska et al., 2010). Mouthing was found to be associated with more frequent production of a wider range of supraglottal consonants in babble. The authors argue that mouthing introduces babies to the feeling of creating closures with the oral articulators somewhere other than at the back of the vocal tract, giving babies incidental exposure to new places of articulation and more enriched “multimodal [auditory and proprioceptive] feedback” (Fagan & Iverson, 2007, pp. 198-199). Similarly, Elbers (1982) observed her own child vocalising with his fingers in his mouth and whilst rhythmically contacting his mouth with the back of his hand between 6-12 months of age and has also suggested that these activities may contribute to the emergence of rhythmic and/or repeated consonant-vowel (CV) syllables. Mouthing may then represent another control parameter, triggering a reorganisation or expansion in the ways that babies can create supraglottal closures. But why, then, does mouthing not cause supraglottal consonants to emerge sooner? By 6 months, babies have typically been mouthing objects for some months (Fagan & Iverson, 2007). However, it is only from around 6 months that babies are reliably able to plan and articulate some adult-like syllables voluntarily (Vihman, 2014). As discussed above, back closures require little precise articulatory control in the first months of life and may be accompanied incidentally by vegetative vocalisation (Kent, 1984; Esling, 2012; Vihman, 2014; Lieberman, 2017). By contrast, moving the tongue tip or the dorsum towards the alveolar ridge or hard palate requires more articulatory control and, when it does occur reflexively (e.g., in nutritive or non-nutritive sucking), it is more rarely accompanied by incidental vocalisation. So then, it may be only after the phonological resources for voluntarily producing CV-like alternations are established that sensory information about new places of articulation gained from mouthing can come to influence what sounds babies may target and acquire next.

Auditory feedback

Finally, Koopmans-Van Beinum et al. (2001) comment on the importance of good quality auditory feedback. Observations that babble may show evidence of sensitivity to the ambient language may illustrate one way in which hearing can influence the form of canonical babble and contribute to later language acquisition (Esling, 2012). Hearing also plays a key role in babies' on-line self-monitoring (Oller & Eilers, 1988; Nathani Iyer & Oller, 2008). Nathani Iyer and Oller (2008) analysed the vocal behaviour of eight hearing, and eight hearing-impaired, babies three months before, at, and three months after babble emergence (termed the *pre-canonical*, *canonical*, and *post-canonical* sessions) and observed some striking differences. Hearing babies were recorded playing with their caregivers or research staff, and hearing-impaired babies were recorded interacting with a Speech and Language Therapist. All adults were requested to encourage babies to vocalise as much as they could. Firstly, the average ages at which these three sessions occurred for hearing and hearing-impaired babies evidenced significant delays in babble in hearing-impaired infants (see Table 1). The authors found no statistically significant difference in the number of utterances and syllables that hearing and hearing-impaired babies produced per minute (their volubility). However, hearing babies produced significantly greater proportions of canonical babble at the canonical session, further increasing between the canonical and post-canonical sessions. By contrast, hearing impaired babies showed a lower proportion of babble during the canonical session, and only afterwards, at the post-canonical session, did they produce "substantial proportions" of babble (Nathani Iyer & Oller, 2008, p. 122). Additionally, hearing impaired babies continued to produce predominantly Fully Resonant Nuclei even after the emergence of babble.

	Hearing babies (yy;mm.dd)	Hearing impaired babies (yy;mm.dd)
Pre-canonical	0;4.0	2;0.0
Canonical	0;7.0	2;3.0
Post-canonical	0;10.0	2;6.0

Table 1 Average age of hearing and hearing impaired infants at pre-canonical, canonical, and post-canonical sessions in Nathani-Iyer and Oller's (2008) study.

Vihman (2014) and Stoel-Gammon (1992) have discussed the contribution to phonological development of a *production/perception feedback loop*. Vihman (2014, p. 109) writes that "[s]ince normally hearing children receive constant feedback from their own vocal productions, they are in a position to accumulate knowledge regarding the aural consequences of phonetic gestures". Commenting on Vihman's (1992) chapter in the same book, in which she identifies well-practised idiosyncratic babble syllables which influence babies' later language acquisition, Stoel-Gammon (1992) discusses how greater volubility and access to this feedback loop can confer significant advantages for hearing children's phonological development. Stoel-Gammon (1992) argues that being more voluble not only means that babies develop more refined neuromuscular control, but it also gives them more opportunities to establish a feedback loop and acquire sensory information from it about how to match auditory signals using their articulatory apparatus – an ability, she notes, that is vital for word acquisition. By contrast, Stoel-Gammon argues, limited or no access to auditory information from this feedback loop may account for the delays in reaching the canonical phase observed in hearing impaired infants (Oller & Eilers, 1988; Stoel-Gammon, 1992; Nathani Iyer & Oller, 2008). Elbers' (2000) chapter on her *Output-as-Input* model discusses the primacy of babies' own vocal output as a source of their internal representations of speech and language. While the chapter focuses on early language acquisition, there are some clear parallels with pre-linguistic phonological development. Elbers (2000) discusses the role of *Double Processing*: when infants vocalise, they provide themselves with haptic, proprioceptive, and auditory-perceptual information about the consequences of their articulatory gestures. This, in itself, is a rich source of phonetic and phonological material but its value may be further enhanced by the effort that is required to process this multi-sensory information. While language acquisition has sometimes been assumed by researchers and theorists to be innate and ‘effortless’, Elbers and Wijnen (1992) argue that effort may in fact play a key role in the transition from babble to first words and in later milestones in language development. They summarise research finding that more explicit and effortful processing typically results in better retention in memory, and argue that information about producing sounds, requiring multimodal processing, should be better retained than information about environmental sounds, which requires only auditory processing (Elbers & Wijnen, 1992).

Elbers (2000, p. 250) goes on to describe language acquisition as “a largely self-constraining process: what is produced in earlier phases sets limits on what can be analysed and acquired, and thereby on what will be produced (and, again, analysed and acquired) in later phases”. This idea of self-constrainment is also explored by Thelen and Smith (1994) and Mareschal et al. (2007) in their discussions of human behavioural and neurophysiological development across all domains. In this sense, hearing and hearing-impaired babies may be constrained in different ways by their experiences with vocalisation. Having different or limited information from which to establish an auditory feedback loop, hearing-impaired babies may be less inclined to repeat resonant vocal sounds that do not provide interesting sensory feedback. This may account for why some hearing-impaired babies begin to produce some canonical babble around the same age as their hearing peers, but do not do this often enough to place them in the canonical phase (Oller & Eilers, 1988). This may also account for delays seen in later language acquisition in hearing-impaired babies (Nathani Iyer & Oller, 2008).

Rhythmic motor development and canonical babble

This section will explore in more depth how rhythmic ability emerges and how it may contribute to phonological development. Between 4-10 weeks of life, babies may begin to produce regular, cyclical, rhythmic movements of the limbs, torso, or the whole body without any clear motive or goal beyond satisfying some endogenous urge (Thelen, 1979, 1981). These behaviours, termed *stereotypies*, are “exceedingly common” in typically developing babies but are believed to occur only atypically in adults, older children, and non-human mammals[4] (Thelen, 1979, p. 699; 1981). As with other behaviours in development, babies show idiosyncratic preferences for different stereotypies; in a longitudinal study of 20 infants, Thelen identified 47 types of stereotypy, and these were often influenced by a baby’s production preferences or their body position (Thelen, 1979, 1981). For instance, rhythmic rocking movements of the torso were most often produced when the baby was seated, or on their hands and knees in a crawl-like position (Thelen, 1979). Stereotypies involving the legs tended to emerge earliest, followed by those involving the torso and the arms or hands. As with other aspects of development, babies’ idiosyncratic preferences and trajectories may give rise to some of the individual differences seen in the order in which different babies attain developmental milestones, and the rate at which they attain them. Different stereotypies involve different muscle systems and may induce different social or biological consequences. For example, arm-flapping may prompt caregivers to supply a baby with a rattle or perhaps a drum to bang. Leg-kicking may strengthen the leg muscles and refine babies’ voluntary innervation of the legs, which may prepare the legs for learning to crawl, stand, and walk later in life (Thelen, 1971, 1981; Thelen & Smith, 1994). Rocking whilst on the hands and knees may give babies some experience with maintaining balance and managing momentum that they may later use in crawling.

The universality of stereotypies suggests that their emergence is determined by endogenous factors like the impulse to engage in exploratory movement or play, and neuromuscular maturation (Thelen, 1979, 1981). Since they are not typically modelled by babies’ caregivers or others around them, it is argued that stereotypies cannot be imitated from adult input (Thelen, 1979). Rather, stereotypies are triggered in the central nervous system once the musculature and neural architecture involved in producing them become functionally mature and before the baby has developed adult-like control over their movements (Thelen, 1981, pp. 3-4). Thelen (1981) has suggested that stereotypies may represent transitional

---

[4] (though healthy birds, reptiles, insects, and fish, however, also produce a high proportion of rhythmic stereotypies (Thelen, 1979, 1981).

phases during which time voluntary control over muscle movement is mastered. Repeated movement strengthens the muscles and the parts of the brain that control them.

In order to affect motoric development in this way, stereotypies must be produced many times (Thelen, 1979, 1981), and something must motivate babies to do this. Producing stereotypies is assumed to have some appealing effect that prompts babies to repeat them. In her 1981 study, Thelen found that babies who received less physical stimulation from caregivers (e.g., rocking, bouncing, carrying) tended to produce more stereotypies, and suggests that these babies' stereotypies may satisfy a biologically pre-programmed drive to stimulate the vestibular system. The vestibular system provides individuals with vital feedback about the location and orientation of the body in in the world and vestibular stimulation is implicated in self-soothing, maintaining balance, and developing neuromuscular control (Thelen, 1981). Individuals who have vestibular impairments or receive insufficient vestibular stimulation can experience problems with balance, strength and agility, vertigo, spatial reasoning, and anxiety (Shumway-Cook, 2007; Carmeli, 2015). The vestibular and neuromuscular systems are underdeveloped at birth in humans and some other primates compared with non-primates (Rosenberg & Trevathan, 2005; Trevathan, 2015). This makes early vestibular development particularly vital in human development. This may partially explain babies' endogenous drive to seek out vestibular stimulation when little is forthcoming. Evolution may have selected for humans to possess some in-built comfort-based imperative or bias towards experiences that result in vestibular development. In the same study, parents reported that their babies produced stereotypies in states of high arousal, i.e., in distress, when fussing or crying, or in excitement upon seeing the caregiver or a source of food. Stereotypies may also, then, serve as a means of releasing affective energy. Piaget (1952, as cited in Thelen, 1979) has also suggested that rhythmic activity benefits cognitive as well as motoric development, since these movements have a perceptible impact on babies themselves and/or their environment, as noted above with reference to arm-flapping, leg-kicking and rocking on the hands and knees (Mareschal et al., 2007). Babies may enjoy the resulting bodily sensations, the sounds produced, or the reactions of their caregivers, and these early multi-modal cause-and-effect experiences may be particularly salient and memorable for babies, requiring the effortful processing involved in the planning and production of a motor movement (c.f. Elbers & Wijnen, 1992). If the consequences of rhythmic stereotypies hold babies' attention in this way, they could provide opportunities to establish a sort of multi-sensory productive/perceptual feedback loop similar to that described for auditory perception earlier in this chapter (Stoel-Gammon, 1992; Vihman, 1992; Vihman, 2014). Indeed, experiments have found that babies below age 1 do have the ability to perceive and attend to the multi-modal consequences of their rhythmic motor and vocal behaviours and will modify their behaviour in response to this feedback (Millar, 1990; Rovee-Collier, 1997; Keren-Portnoy et al., 2021). A further evolutionary and developmental advantage of rhythmic stereotypies may be that the action pattern for producing rhythmic cycles of movement can be "pre-programm[ed]" to some extent in the central nervous system (Thelen, 1981, p.3). In the brains and body plans of less complex, differentiated organisms like insects or birds, this may allow, for example, for rhythmic sounds to be produced simultaneously with other actions like nest-building that require more flexible, context-dependent control (Vallortigara et al., 1999). In humans, pre-programmed, automatised stereotypies may have the effect of 'freeing up' cognitive or motoric resources, which may then be used to carry out more complex, context-dependent activities simultaneously (Thelen & Smith, 1993; Vihman, 2014). Rhythmic stereotypies may also play some role in the destabilisation and reorganisation of behavioural systems, including babble.

The coupling of the rhythmic and vocal systems

Around the time that babble emerges, babies also begin to show developments in generalised rhythmic motor activities (Thelen, 1979, 1981; Ramsay, 1980, 1984; Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001; Iverson & Fagan, 2004; Iverson et al., 2007). Empirical and observational evidence suggest that the rhythmic and vocal systems are closely linked and that developmental advances in each system influence those in the other. Like Thelen (1979, 1981), Iverson and Fagan (2004) note that babies produce rhythmic motor movements for some time prior to the onset of canonical babble, but some research has identified an increase in rhythmic behaviours at around 24-32 weeks (i.e., 6-8 months) – around the time when reduplicated babble typically emerges (Thelen, 1979, 1981; Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001; Iverson & Fagan, 2004; Iverson et al., 2007). In Ejiri and Masataka's (2001) longitudinal study, babies produced significantly more rhythmic movements (like repeated leg-kicking or arm-flapping) than non-rhythmic movements (like handling and mouthing) whilst vocalising pre-canonically. A peak in simultaneous vocalisation and rhythmic movement was identified during the month before babies entered the canonical phase (between 5-7 months of age in this study). This rhythmic body movement later declined “noticeably” in the months after each baby reached the canonical phase, while canonical syllable production continued (Ejiri & Masataka, 2001, p. 43). This decline has not been replicated in other similar research, however, and this may reflect idiosyncratic production preferences amongst their relatively small sample of four babies. Iverson and Fagan’s (2004) longitudinal study, which involved a larger sample of 47 infants, found that 20% of rhythmic movements were coupled with vocalisation and that this *increased* slightly over time between 6 and 9 months of age. Iverson and Fagan (2004) also remark on some methodological differences between their criteria for identifying overlap between vocal and generalised rhythmic behaviours and those used by Ejiri and Masataka (2001).

A particularly strong relationship has been identified between babble and rhythmic *arm and hand* movement specifically (Ejiri, 1998; Iverson & Thelen, 1999; Iverson et al., 2007). Babies typically begin to simultaneously produce babble and rhythmic manual movements at around 6-8 months of age (Iverson & Thelen, 1999). Of the 20% of rhythmic motor movements that occur simultaneously with vocalisation in Iverson and Fagan’s (2004) study, the majority of these were movements of the hands rather than movements of other parts of the body. In a longitudinal study of 26 babies, Iverson et al. (2007) found a near fourfold increase in rhythmic arm movement, correlated with infants’ arrival at the canonical stage, but not with their chronological age, since babies in the study began to produce canonical babble anywhere between 4 and 9 months of age. The relationship between rhythmic development in the manual and vocal systems cannot, therefore, be attributed simply to maturation but likely also involves some experience-dependent process of mutual entrainment. Since rhythmic motor movement precedes, and is later co-produced with, rhythmic vocalisation, the properties of rhythmic hand and arm movement may offer some insight into how the rhythmic temporal organisation of babble emerges. As discussed above, rhythmic ability may represent an established, stable capacity that ‘attracts’ a relatively new and unstable form – marginal babble – and triggers a reorganisation in it. When a baby vocalises whilst moving their hands rhythmically, rhythmicity may ‘spill over’ from the more stable behaviour – the stereotypy – and influence the way that the baby produces the less stable behaviour. This may come about through the spreading of neural activation in a *synergy* or network of brain areas, responsible for producing complex co-ordinated movements. Research has found evidence for the existence of such a synergy in human adults and infants as well as in macaques, suggesting that interconnectedness between the hand and mouth is a phylogenetically ancient trait (Gentilucci & Corballis, 2006; Petrides & Pandya, 2009; Gentilucci et al., 2012; Desmurget et al., 2014).

It has also been posited that simultaneous production of rhythmic hand and arm movement with canonical babble may be a precursor to simultaneous speech and gesture production. Iverson and Thelen (1999) note that users of all spoken languages can be observed to produce speech and gesture simultaneously, even in cases where both or all interlocutors are blind, suggesting that this is not a tendency that we acquire from observing and re-enacting the behaviour that we observe from those around us (c.f. Donald, 1991 on mimesis). The authors argue, instead, that the coupling of oral and manual motor movement may be innately pre-specified in our neurophysiology, and they cite the Babkin reflex – the automatic (or reflexive) opening of the mouth and flexion of the arms towards the face when pressure is applied to the palm/s of newborn infants (Futagi et al., 2013) – in support of their argument. When newborn babies mouth their own hands or objects that they are holding in their hands, their mouths open anticipatorily before the hand or object reaches the mouth, suggesting that there is some simultaneous neural activation in the manual and oral motor areas at a time well before babies begin to show evidence of the capacity for coordinated motor planning (Iverson & Thelen, 1999).

Desmurget et al. (2014) suggest that the hand/mouth motor synergy that they identified between the hand and mouth motor cortical areas in the precentral gyri may be established before birth and may account for babies’ and foetus’ ability to raise the hand/s to the mouth with perhaps surprising accuracy, while their oral and manual neuromuscular control and co-ordination is otherwise immature and imprecise. These motor areas may never become *de*coupled, due their frequent co-activation from at least the time of birth and perhaps also during foetal movements *in utero*. Gentilucci and Corballis (2006) observed that handling different sized objects or producing manual gestures of different sizes causes differences in the degree to which adults open their mouths when producing vowels, and the same happens when an adult simply watches another human gesturing or handling objects. The hand-mouth coupling may be constantly reactivated and thus repeatedly myelinated[5] during our everyday interactions with our environment and those in it. The neural pathways connecting the manual and oral motor areas may be resistant to synaptic pruning because they are so frequently reactivated – and so often strengthened – when we produce simultaneous oral and manual movements throughout infancy childhood, and adulthood (Iverson & Thelen, 1999, p. 20).

The temporal organisation of rhythmic stereotypies and babble

As discussed in the previous section, emergent rhythmic capacity plays a critical role in babble and, thus, in the ontogeny of speech and language, representing the control parameter that triggers the transition from arrhythmic marginal babble into canonical babble, which is characterised by more regular rhythm. In this next section, further evidence for this claim will be discussed.

Research by Meier et al. (1997) found evidence that babies' vocalisations may be sensitive to rhythmic movements of the lower jaw. Some hearing and hearing-impaired infants who had begun to produce canonical syllables were also observed to produce unvoiced rhythmic openings and closures of the jaw, termed *jaw wags*. These silent movements were produced sometimes in isolation, and sometimes in sequences which also contained canonical babble. Since jaw wagging, involves movement of the lower mandible – typically the primary articulator in babies' earliest babbles – it seems intuitive (taking a view of development as a self-determining and self-constraining process), that the rhythmic properties of jaw wagging should come to influence the way that these babies form sequential openings and closures in the vocal tract during canonical syllable production. However, in a sample of 14 babies, all of whom were observed to produce canonical babble, Meier et al. (1997) observed jaw wagging in only seven. Jaw wagging cannot, therefore, be taken to represent the primary or only source of the rhythmicity observed in canonical babble. Rather, rhythmic movements in seemingly more geographically distant parts of the

---

[5] See section *Broca's area as a "precise timing mechanism"* later in this chapter for further explanation

body may also contribute to the rhythmic reorganisation of babies' vocal behaviour, mediated by motor production synergies in the precentral gyrus (Desmurget et al., 2014). Further to the findings discussed above, acoustic analysis has revealed that infant vocalisations produced simultaneously with rhythmic whole-body movements show short formant frequency transitions and syllable durations, comparable to those found in adult speech (Ejiri & Masataka, 2001). By contract, infant vocalisations with no accompanying rhythmic movement showed longer syllable durations and slower formant frequency transitions (Ejiri & Masataka, 2001). Mature formant transitions are amongst the "infraphonological", "metaphonological", or "subphonological" features, which characterise babble and which are not present in earlier vocal behaviour (Oller, 1980, p. 93; Ejiri & Masataka, 2001, p. 47; Nathani Iyer & Oller, 2008, p. 322).

Ejiri and Masataka's (2001) findings constitute robust evidence of a one-way interaction or *entrainment relationship*, with rhythmic whole-body motor movement influencing the way that babies produce vocalisations. Other research suggests that a *mutual* entrainment relationship may exist between vocal and motoric systems. Iverson and Fagan (2004) found that babies who had started to babble showed more than twice as much simultaneous vocalisation and rhythmic *hand* movement (rather than whole body movement) as babies who had not yet begun to babble. This hand movement may also now be unimanual, in contrast with earlier bimanual rhythmic movement. So then, the stable, regular rhythmicity of babies' stereotypic movements may act as a sort of 'magnet', attracting marginal babble and influencing the less regular and less precise movements that babies make when producing it. The shorter duration of each cyclical motion in a rhythmic stereotypy limits the time available for moving the articulators between opening and closure, and this results in shorter formant frequency transitions between the consonantal and vocalic portions of canonical syllables. This does, of course, rely on the baby having established the neuromuscular control required to move the articulators sufficiently quickly, and the ability to plan motor actions, as discussed earlier in this chapter. In turn canonical babble may prompt babies to produce more rhythmic manual behaviour. When the rhythmic capacity and the mouth areas in the motor cortex are co-activated, neural activation may spread through integrated synergies to other connected parts of the brain, like the hand areas of the motor cortex. Through repeated and spreading co-activation, behaviours and capacities can become *time-locked* or bound together through instances of co-production (Thelen & Smith, 1994).

Broca’s area as a “precise timing mechanism”

Iverson and Thelen (1999, p. 22) suggest that vocal and manual motor behaviour may share a “common precise timing mechanism” housed in the lateral perisylvian cortex – the parts of the brain surrounding the Sylvian fissure or lateral sulcus, which is the deep, easily identifiable groove where the frontal and temporal lobes meet, and which extends into the parietal lobe. This area comprises the secondary somatosensory cortex (Brodmann areas 40 and 43), Wernicke’s areas (Brodmann area 22), and Broca’s area (Brodmann areas 44 and 45 in the inferior frontal gyrus in the left[6] hemisphere) (Mareschal et al., 2007).

It is relevant, for the purposes of this thesis, to note that all complex behaviours are underlain by activity in networks of neurons that are situated in different regions of the brain, and that may be involved in several different functions. Neural structures that are responsible for generalised motor actions are also involved in speech and language production in adult humans, and structures often thought of as ‘linguistic’ are also activated in other, non-linguistic behaviours (Iverson & Thelen, 1999; Iverson et al., 2007). For example, neural connections have been identified between parts of the cerebellum – classically associated with motor activity – and parts of the cerebral cortex that are involved in speech and language perception and processing like Brodmann area 8 (in the superior prefrontal cortex) (Leiner et al., 1993). The ventral dentate nucleus in the cerebellum has even been found to show activation during tasks that draw on cognitive and linguistic capacities but involve no motor movement at all, indicating a critical role for the cerebellum in speech and language perception as well as production (Leiner et al., 1993). The dentate nucleus is a tooth-shaped area located at the centre of the cerebellum in each hemisphere (Leiner et al., 1993). The dorsal dentate nucleus projects to the primary motor and premotor cortices (Brodmann areas 4 and part of Brodmann area 6 respectively), while the ventral dentate nucleus projects to areas in the parietal lobe (involved in sensory perception) and

---

[6] Broca’s area is located in the dominant brain hemisphere. This is typically, but not exclusively, the left hemisphere.

prefrontal areas (involved in attention, memory, impulse control, and flexible cognition) (Dum & Strick, 2003).

By contrast, and of particular relevance for this chapter, part of Broca's area, typically associated with speech and language processing, has been implicated in non-speech motor functions in humans and non-humans (e.g., Iverson & Thelen, 1999; Cantalupo & Hopkins, 2001; Hecht & Parr, 2015). The pars opercularis (Brodmann area 44 in humans) is activated when we plan and produce manual motor behaviours, and when we watch the manual activity of others with the goal of interpreting their behaviour (Leiner et al., 1993; Fadiga & Craighero, 2006; Skipper et al., 2007). Iverson and Thelen's (1999, p.22) argument for Broca's area as such a "common precise timing mechanism" aligns well with the evidence from neuroimaging and neurosurgical research. Whilst the evidence they cite for this argument is largely drawn from adult neurophysiology, Iverson et al. (2007) have argued that there is reason to believe that these neural connections may exist from birth and may be strengthened through continued reactivation over the lifespan, as discussed previously in this chapter. Iverson and Thelen (1999) further suggest that the coupling of the vocal and manual motor systems seen in babble and in speech may make use of ancient and fortuitous neural connections that have formed during earlier stages of human evolution through production of other useful functions. Recall that speech sound production makes use of physiological structures that initially evolved for purposes unrelated to speech, language, or even sound production like the larynx, jaw, or tongue (Koopmans-Van Beinum et al., 2001; Esling, 2012). The evidence discussed briefly in this section suggests that the brain regions and mediating pathways between them involved in speech, language, and babble also evolved to serve general non-linguistic functions (Leiner et al., 1993; Iverson & Thelen, 1999; Dum & Strick, 2003; Fadiga & Craighero, 2006; Iverson et al., 2007; Skipper et al., 2007). This perspective and other related research will be discussed in more depth in Chapter 3 of this thesis.

In the previous section of this chapter, it was argued that the mutual entrainment relationship between babble and rhythmic whole-body and manual movement may be explained in terms of spreading activation between the mutually connected or *synergistically integrated* neural sites responsible for rhythmic ability and innervation of the hands and mouth. In light of evidence that Broca's area is implicated in spontaneous and imitated oral and manual motor articulation and planning, speech and language, and rhythmic and sequential coordination, it seems that developments in Broca's area may play some role in the emergence of babble and in the coordination of simultaneous rhythmic vocal and manual behaviours.

Repeated co-activation of Broca's area and the hand and mouth motor areas may strengthen the neural pathways between these areas. Neural pathways are strengthened through *myelination* (Mareschal et al., 2007). When neural pathways are activated, the *axons* (cable-like projections) of each neuron in a network become wrapped in a segmented sheath made of myelin: a fatty, proteinous substance (Williamson & Lyons, 2018). These myelin sheaths act as protective electrical insulators and allow electrical signals to pass through the axons of neurons more quickly and efficiently (Su et al., 2008; Williamson & Lyons, 2018).

Whilst myelination begins *in utero*, the newborn brain contains little myelin, and this is restricted to sensory processing areas, while myelination in other parts of the brain continues to develop rapidly until age 18-36 months and at a slower rate thereafter (Hayakawa et al., 1991; Su et al., 2008; Rosselli et al., 2014; Williamson & Lyons, 2018). With regard to Broca's area specifically, research by Su et al. (2008) found no evidence of myelination in newborns, and that Broca's area (along with Wernicke's area and the angular gyrus) showed slower myelination than the motor, auditory and visual cortices, but faster myelination than the arcuate fasciculus (a bow- or arc-shaped bundle of axons connecting Broca's and Wernicke's areas). Experience-dependent strengthening of neural pathways through myelination plays a critical role in general brain development, and so supports our mastery of motoric, cognitive, and linguistic skills, amongst others (Williamson & Lyons, 2018). Babies in Su et al.'s (2008, p. 1761) study showed mature myelination in the "language-correlated regions" under investigation by 18 months of age and the authors suggest that this may partly underlie the acceleration typically seen in vocabulary growth at around this age. Myelination may, by 18 months, serve to make word processing efficient enough to 'free up' sufficient attentional, productive, and cognitive resources to acquire and process new words and concepts. While 18 months is typically long after the time that babies begin to babble, most *major* changes in myelination occur within the first 8-12 months of life (Van Der Knaap & Valk, 1990; Grotheer et al., 2022) meaning that, whilst still immature by the time of babble emergence at 6-8 months, myelination has very likely already begun in Broca's area by this time, and babble production may result from and contribute to its continued strengthening.

The Babkin reflex and its contribution to babble emergence

Futagi et al. (2013) suggest that the Babkin reflex may originate in the shared phylogeny of humans and some other higher mammals (like primates and rodents) and may support feeding. The Babkin reflex is implicated in breastfeeding in a similar way to the plantar reflex in other animals[7]: during suckling, oral stimulation prompts babies to use the palm to increase milk flow from the breast (Vallone & Carnegie-Hargreaves, 2016) and the reflex is suppressed in the period immediately after feeding (Futagi et al., 2013). However, Futagi et al. (2013) have pointed out that the reflex may not be wholly helpful during breastfeeding, since stimulating the palm can cause the mouth to open too widely to latch onto the nipple. The reflex may also play some role in supporting the development of *self*-feeding. The automaticity of Babkin reflex is partly underlain by activity in the reticular formation in the brainstem, adjacent to the cerebellum (Futagi et al., 2013) and producing the reflex involves simultaneous innervation of the hand and mouth motor cortical areas. The Babkin reflex may serve to 'protect' the biologically pre-programmed neural connections between the hand and mouth areas from synaptic pruning through repeated activation until such a time as babies are able to innervate these areas voluntarily. By around the fifth month of life, the Babkin reflex typically declines and is replaced by less automatic, more flexible and adaptive hand-to-mouth movements as babies develop the neuromuscular control to bring objects to the mouth voluntarily for exploration or feeding (Futagi et al., 2013). With time and repeated experience, these voluntary hand-to-mouth movements become more synchronised, fluent, and adult-like. So then, the Babkin reflex may represent another mammalian capacity that humans make use of in phonological development. A mechanism that has evolved to support the development of nutritive behaviour may also shape babies' exploration of the vocal tract through mouthing the hands and other objects (Fagan & Iverson, 2007) and may serve to protect the neural connections between the hands and

---

[7] The plantar reflex is commonly observed in kittens during feeding and older cats during bedding and is often colloquially described as "making biscuits" or "kneading bread".

mouth, which later become associated with the rhythmic, motoric, and cognitive capacities housed in Broca's area.

Qualitative changes or milestones in canonical babble

The remainder of this chapter will consider phonological milestones that babies attain following the emergence of canonical babble and their significance for the transition into first word acquisition.

Singleton, reduplicated, and variegated babble

As we have seen earlier in this chapter, canonical babble denotes three distinct types of infant vocalisations, which share certain concrete phonological properties but also index the attainment of new skills. Singleton babble indicates the establishment of more adult-like articulatory skills, and reduplicated babble is associated with rhythmic ability, while variegated babble shows evidence of more sophisticated articulatory planning. There is some degree of disagreement amongst authors as to precisely when babies arrive at these milestones (Oller, 1980; Stark, 1980; Studder-Kennedy, 1990; Vihman, 2014). This may reflect the great variability seen between the developmental trajectories of individual babies and young children (Oller, 1980; Vihman 2014). For instance, the child in Elbers' (1982) diary study produced reduplicated babble before singleton babble, and this may have been an idiosyncrasy, influenced by the other behaviours and capacities that this child had established in repertoire at the time of babble emergence. it is possible that babies who begin to babble very early may produce singleton babble first, while babies who begin to babble later may have a more developed capacity for rhythmic movement by the time that their first syllables emerge and so may produce reduplicated babble first. Recall that, while the emergence of canonical babble is universally observed across all developing babies (Oller, 1980; Stark, 1980; Studdert-Kennedy, 1990, Vihman, 2014), Oller (1980) does provide the caveat that these non-discrete phases within babble represent *tendencies,* which may not be universal or universally ordered. It may be that the extent of individual variation in babies' developmental trajectories and the highly transitional nature of these phases in phonological development make it difficult to determine exactly when singleton and reduplicated babble emerge in relation to one another.

What is more widely accepted is that variegated babble emerges some time after reduplicated or singleton babble (Oller, 1980; Stark, 1980; Elbers, 1982; Oller, 2000) and that variegated babble therefore represents a more sophisticated behaviour. Mitchell and Kent (1990), however, have claimed that babies produce variegated babble around the time when babble first emerges, and that variability in consonant production may even decline with age and experience. In a longitudinal study of four babies, they found no statistically significant overall difference in the extent of variegation over time, with only a slight u-shaped regression in variegation. Only one baby showed a significant increase in variegated babble between 0;7.0, 0;9.0 and 0;11.0, and one baby showed a significant decline. Mitchell and Kent's (1990) conclusion is based on a small and diverse sample, so may be vulnerable to the effects of individual variation. While singleton and reduplicated babble require the ability to plan (and repeat) one supraglottal closure and opening, variegated babble requires the ability to plan two or more closures in distinct parts of the vocal tract. Varying consonants voluntarily therefore poses a higher articulatory and cognitive load than producing (and repeating) a single consonant. However, working once again on the assumption that development is self-shaping (Elbers, 2000; Mareschal et al., 2007), this cognitive and articulatory load may be reduced if some part of the motor action pattern involved is already well-practiced and established in repertoire. For instance, producing [daba] may require less effortful processing if a baby is already proficient at producing [dada] or [baba]. Experience of sequential production may be even more valuable for much longer variegated sequences that are also produced by some babies like the utterance [utɪgəʃəw], produced by Leif, aged 0;11.23 in the present study. Although vowels and manner of articulation are usually not considered when analysing babies' developing phonological systems, this example from Leif is still highly complex, containing consonants with four distinct places of articulation. Leif began to produce canonical syllables at 0;5.30, though he did not enter the canonical phase until 0;8.15. At both points in time, he showed a strong preference for labial articulations and also produced some coronals at the latter time. At 0;8.15, Leif was also producing reduplicated babble and continued to do so throughout the data collection period. His early and continuing experience of reduplicated babble with labials may have afforded him enough entrenched productive experience to act as a ‘springboard’, from which he could begin to expand his consonant repertoire and combine consonants into babble sequences by 0;11.23. As babies become more proficient babblers, their phonological skills begin to diversify. In the following sections, we will see how other entrenched capacities also act to decrease effort in phonological perception and production, ‘freeing up’ resources to acquire new, more complex skills, resulting in cyclical reorganisations of the systems and subsystems underlying phonological development.

Vocal Motor Schemes

Another phonological milestone within babble that is of particular significance for later word acquisition and for this thesis is the emergence of favourite consonants or *Vocal Motor Schemes*. These are well-practiced idiosyncratic production routines that allow babies to produce certain consonants as “consistent phonetic forms” (McCune & Vihman, 2001, p. 673; Vihman, 2014). McCune and Vihman (1987) coined the term after Piaget’s (1962) *action schemes*, which Piaget (1952, p. 7) describes as “cohesive, repeatable action sequence[s]” formed from “component actions that are tightly interconnected”. In Dynamic Systems terms, this can be understood as an interaction of component behaviours, time-locked together, with each constituting a subsystem of an overlying system or behaviour. In this case, that overlying behaviour is the ability to produce a given consonant consistently. Vocal Motor Schemes typically emerge when babies are around 9-12 months of age, but sometimes later, as their emergence is driven by articulatory experience. Vocal Motor Schemes become entrenched and stable through repeated production between 12-18 months of age, and represent familiar routines by which babies can reliably produce a particular consonant or consonants (McCune & Vihman, 1987, 2001; Vihman, 2014)

Vocal Motor Schemes are identified as follows. Longitudinal naturalistic observational data is collected in a familiar place wherever possible (like the home of the family or a friend), or, in some cases, in a laboratory setting at intervals of no longer than one month (McCune & Vihman, 1987, 2001; Vihman, 2014). Infant vocalisations in this data are transcribed using the International Phonetic Alphabet and the incidence of consonant production is analysed. If the resulting transcripts from a series of >30 minute recordings contain either ≥10 productions of a consonant in each of three out of four consecutive sessions, ≥30 productions in one single session, or ≥50 productions over one to three sessions, then this is taken to indicate that that baby has a well-established production routine for that consonant (McCune & Vihman, 1987; McCune & Vihman, 2001; Keren-Portnoy et al., 2010; DePaolis et al., 2011; Vihman et al., 2014). Voicing contrasts and vowels are not considered since there is little evidence that infants have voluntary control over voice onset times or mastery of the vowel space before 18 months of age (Macken & Barton, 1980; Stark, 1980; DePaolis et al., 2013).

The origins of Vocal Motor Schemes

It is as yet unclear precisely how each baby comes to settle on their Vocal Motor Scheme/s (DePaolis et al., 2011). One psychologically and physiologically plausible explanation may be found in the principle of "least effort" (Elbers, 1982, p. 57). Labials and coronals are very common Vocal Motor Schemes (Vihman & McCune, 1987; McCune & Vihman, 2001; DePaolis et al., 2011, 2013; Majorano et al., 2014; McGillion et al., 2017). As discussed earlier in this chapter, these sounds can be produced using a relatively simple raising and lowering of the mandible and so these may represent easier consonants for babies to produce while their neuromuscular control over their vocal apparatus is still immature (MacNeilage, 1998; Nip et al., 2009). Neuromuscular control over the jaw matures some time before control over the tongue and lips: babies are able to produce adult-like jaw oscillations by around 12 months of age, but adult-like control over the lips is not seen until 2-6 years of age (Nip et al., 2009). Kent (2021) reviews research finding that, while movements of the tongue become quite sophisticated through sucking by around 14 weeks of age (p. 1584), the movement of the lips is controlled primarily by muscles in the cheeks and jaw until some time later in development. Kent (2021, p. 1591) offers the following explanation: "It is understandable that motor control of the mandible in children is developed before that of the lips. Otherwise, precise control of the lips would be jeopardized by inaccuracies in jaw movements." The ease of articulation argument has been cited in support of Davis and MacNeilage (1995) *Frame/Content theory* regarding the articulatory basis of babble: the *frame* portion (the closed-to-open movement) of syllables containing alveolar and labial consonants may be easier for babies to produce than syllables containing approximants or palatal, velar, or pharyngeal stops, which requires simultaneous jaw oscillations and finer motor control over the tongue (Locke, 1983; MacNeilage, 1998). Frame/Content theory will be discussed in more detail in Chapter 2.

Another possibility, and one which would contradict maturational accounts of babble and language acquisition, including the Motoric hypothesis, is that Vocal Motor Schemes may be influenced by the sounds that babies hear in the speech of adults and older children around them. Amongst babies whose ambient language is English, [t/d] is most commonly favoured, whilst babies whose ambient language is Italian tend to prefer [p/b] (McCune & Vihman, 2001; DePaolis et al., 2011; DePaolis et al., 2013). In their study of 53 infants acquiring British English (n=27) and Welsh (n=26), DePaolis et al. (2013, p. 647) remark that alveolar stops are "motorically accessible and abundantly modelled in the input" and suggest that the interaction of these factors may account for the prevalence of [t/d] as a Vocal Motor Scheme amongst babies in their sample. The same reasoning may account for the prevalence of [p/b] in the babble of Italian babies (Majorano et al., 2014). However, the relationship between ambient language and babies preferred consonants does not seem to be mediated straightforwardly by input frequency. While some babies have been found to favour consonants that *are* very frequent in the language to which they are exposed, some other infants have been found to favour the consonants that are amongst the *least* frequent consonants in the input (Oller, 1976; Vihman et al., 1985). In fact, Elbers and Ton (1985) have found that the linguistic input does not even reliably determine babies' production preferences in their first words. The infant in their diary study, Thomas, showed an idiosyncratic production preference for utterances of the form /$C_{labial}VC_{alveolar/}$; a pattern which is infrequent in the child's ambient language, Dutch. If input frequency cannot reliably determine the phonetic form of children's first words (a behaviour which is overtly sensitive to the linguistic input) this weakens the claim that production preferences in babble (as a behaviour that is further removed from speech and language) should be influenced by language exposure. Many world languages make use of fricatives, yet babies' babble relatively rarely does (Oller et al., 1976; DePaolis et al., 2011). Creating the stricture for producing fricatives intentionally and reliably requires a higher degree of articulatory precision than for producing stops. Babies in the present study tended strongly to prefer stops and showed a marked preference for [t/d] (see Table _ for further detail). Whilst some of these babies did develop fricated and approximated Vocal Motor Schemes, these tended to emerge after babies already had at least one stop in repertoire. These observations, along with the findings of cross-linguistic studies of babble are better accounted for by the ease of articulation argument.

Again, Vocal Motor Schemes are highly idiosyncratic: while [t/d] may be a particularly commonly favoured consonant amongst English-acquiring babies, this finding reflects a general tendency over large sample sizes of babies (McCune & Vihman, 2001; DePaolis et al., 2011; DePaolis et al., 2013). Babies show much individual variation both in terms of which consonants they favour and in terms of how many Vocal Motor schemes they develop (Vihman, 2014). This is reflected in the findings of the present study. Whilst 10 out of 12 babies have [t/d] as a Vocal Motor Schemes, they show much variation in their other Vocal Motor Schemes. Several babies (Cameron, Fred, Leif, Maebh, and Morgan) developed only one Vocal Motor Scheme before 12 months of age, whilst Bella produced seven consonants frequently enough to qualify as Vocal Motor Schemes. It seems, then, that Vocal Motor Schemes may be ‘selected’ or determined according to endogenous factors. Differences in the morphology of the vocal tract like, for example, a slightly larger tongue or a shallower alveolar ridge may make it more or less likely that particular babies may happen upon particular sounds during exploratory play. Differences in the neural connections that are in place from birth may similarly bias some babies to produce certain sounds more or less often. As discussed, articulatory experience is cumulative and self-constraining, and producing a particular sound may prompt a baby to produce that sound again if the sound or feeling was enjoyable or if it satisfied the baby’s drive to behave like those around them (Donald, 1991; Vihman, 2014). Babies may ‘discover’ sounds accidentally in the course of their ’haphazard’ exploration of their own vocal apparatus, complemented by exploratory mouthing (Fagan & Iverson, 2007) and may come to fixate on particular sounds because they are enjoyable or easy to produce again and again. Over time, babies’ developing physiology, perceptual awareness, and articulatory control may become coupled together along with newly emerging capacities for memory and attention, motor planning, and rhythm through time- and experience-dependent interactions as the baby engages with their environment and their own body (Stark, 1980; Thelen & Smith, 1994; Davis & MacNeilage, 1995; MacNeilage, 1998; Mareschal et al., 2007; Esling, 2012; Vihman, 2014). These couplings may give rise to the emergence of favoured sound production routines, which then become consolidated or ‘entrenched’ through repetition over multiple instances.

Several studies have found evidence strongly supporting the idea that the production-perception feedback loop (Stoel-Gammon, 1992; Vihman, 1992), rather than linguistic input, is instrumental in strengthening babies’ representation of, and preference for their Vocal Motor Scheme/s. Whilst infant-directed speech across caregivers exhibits common phonological and phonotactic properties, babies’ own production preferences show significant interindividual variability and significant deviation from the speech to which they are directly exposed (Vihman et al., 1994; DePaolis et al., 2011; DePaolis et al., 2013). Elbers and Wijnen (1992, p. 339) comment that “what is processed with effort tends to be well retained”. Production, being more effortful than perception, may leave a stronger trace in babies’ phonological memory, meaning that self-produced sounds are better represented and remembered than sounds perceived in the linguistic input. Similarly, Elbers’ (2000, p. 250) Output-as-Input model views babies’ own production as a “privileged” source of input because it provides auditory-perceptual and proprioceptive feedback.

Vocal Motor Schemes and Waddington’s epigenetic landscape model

The emergence and entrenchment of Vocal Motor Schemes can be understood with reference to Waddington’s (1966, 1977) *epigenetic landscape* analogy (cited in Thelen & Smith, 1994). In this analogy, development is represented by a rolling ball that moves between attractor states or *wells*. These wells vary in depth according to the stability of the behaviour that they represent: more stable behaviours, like canonical babble, are represented by deeper wells and less stable behaviours, like marginal babble, are represented by shallower wells. Whilst in a well, the rolling ball may be subject to *perturbation* in real time and over developmental time. Real time perturbations may be understood as small ‘bumps’ that move the ball around within or between wells and may include things like single instances when a baby discovers a new sound, perhaps through mouthing. Over developmental time, the cumulative effect of many real time perturbations, or repetitions of a single perturbation, may cause the walls of a well to weaken and crumble, thus lowering the threshold that must be crossed in order to enter or leave a well (Iverson & Thelen, 1999). This increases the likelihood that the ball may be bumped out of or into a well each time it is perturbed. These thresholds may also be lowered or overcome by single instances of very strong perturbation. In Dynamic Systems theory terms, perturbations that cause the ball to exit a well may be understood as *control parameters* and movement between wells can be understood as *system reorganisation.* Shallower wells are more vulnerable to outside influence in real time, so behaviours with shallower wells may only be produced for a relatively short period of time. Deeper wells are more immune to real time perturbation and the rolling ball may rest in these for much longer.

Behaviours with deeper wells are also more flexible; once firmly established through repetition, a behaviour may be produced in different contexts without the ball being prompted to roll out of the well, that is, without changing the form of the behaviour (Waddington, 1966, 1977; Thelen & Smith, 1994). Experience can affect the depth of these wells (Thelen & Smith, 1994). When a baby discovers a sound that they enjoy producing and that they are motivated to repeat frequently, a deep well may form, in which the rolling ball may come to rest for many weeks or months. At first, while a well is forming and its threshold is being traversed, this sound may be effortful to produce and may require some attention. Once well-embedded and well-practiced, this sound becomes available in repertoire and can be produced flexibly and with less concentration and effort and less likelihood of change in the form of the sound. A baby who enjoys producing raspberries may first produce them in isolation with some effort and may later produce them with less effort whilst eating or vocalising – or both at the same time! Similarly, a baby who enjoys producing labial sounds may first produce them in isolation, then in their stable babble and, later, once a deep well has formed, may begin to make use of these sounds in their first communicative utterances when the cognitive capacities for intentional communication and symbolic representation emerge (McCune & Vihman, 1987). This analogy may also account for why babies who have more than one Vocal Motor Scheme can seem to have stronger and weaker consonant production preferences. The depth of these wells being a gradient feature, Vocal Motor Schemes that are more frequently produced in babble or that exert more influence over word form targeting and production may have deeper wells than others.

The role of Vocal Motor Schemes in phonological development

Vocal Motor Schemes were initially proposed by McCune and Vihman (1987) in a study seeking to investigate the degree of consistency between consonant production in babble and that in babies' first words. The authors noted that babies in their study had preferred consonants that appeared in their intentional vocalisations and could be combined with other phonetic material to form larger units, and that these consonants tended to feature prominently in their first words. Marked similarities have also been recorded between babies' production preferences in their babble and in their first meaningful word forms in other research before and since this study (Oller et al., 1976; Elbers, 1982; Locke, 1983; Elbers & Ton, 1985; Oller & Eilers, 1988; Elbers & Wijnen, 1992; Vihman, 1992; Vihman, 1993; Vihman et al., 1994; Elbers, 2000; Oller, 2000; McCune & Vihman, 2001; Keren-Portnoy et al., 2005; Keren-Portnoy et al., 2010; DePaolis et al., 2011; DePaolis et al., 2013). The findings from these and other studies constitute further strong evidence that babble does, indeed, represent a dynamic and self-shaping system.

Vihman et al. (2009, p.210) describe Vocal Motor Schemes as the origin of the "raw material" for word acquisition, since mastering voluntary control over the vocal tract is a necessary precursor to word acquisition. In Dynamic Systems terms, Vocal Motor Schemes may themselves act as control parameters, motivating a synergistic reorganisation of generalised vocal, motor, and rhythmic capacities, along with perceptual and proprioceptive sensory information gained from using these capacities. Before they are able to produce consonants voluntarily, babies are ill-equipped for language (Vihman, 2014). Some babies may develop the cognitive capacity to *understand* some of the arbitrary sound-meaning relationships of language before Vocal Motor Scheme emergence, but any attempts at word production may be limited, effortful, or inconsistent. However, once babies can produce at least one consonant proficiently, with relatively little effort, this 'frees up' articulatory and cognitive resources, which can then be used for functions like applying known production routines to new contexts, processing symbolism, and processing others' social cues about attention, intention, and emotional state (Meins, 1998; Munday et al., 2000; Vihman, 2014; McGillion et al., 2017).

Vocal Motor Scheme emergence marks a significant qualitative change in the form of babble and in when and how babies produce it, allowing babies to behave more and more like those around them. As with any other developing behaviour or skill, this growth in articulatory proficiency is likely to be underlain by some change in the neural architecture that supports it. This idea will be discussed further in Chapter 3.

The articulatory filter

The idea of the *articulatory filter* was proposed by Vihman to explain how and why babies' first words tend to be relatively accurate renditions of the target form given their limited phonological systems (Vihman, 1993; Vihman et al., 2014). Vihman (1993, p. 74) defines the articulatory filter as "a phonetic template (unique to each child) which renders similar patterns [to their own Vocal Motor Schemes] in adult speech unusually salient or memorable". The salient sounds or patterns of sounds act as an "entry-point", giving babies landmarks by which they can begin to segment the speech stream (Vihman, 1993, p. 73). The theory is supported by empirical and observational evidence from a number of studies (Keren-Portnoy et al., 2005; Keren-Portnoy et al., 2010; DePaolis et al., 2011; DePaolis et al., 2013; Vihman, 2014; Vihman et al., 2014). Whilst babies' first words are said to be accurate or faithful to the adult target form, that is not to say that they are mature, adult-like forms with wholly target-like underlying representations. Rather, they are globally accurate holistic representations of word forms, which mediate between the adult target form and the perceptual, cognitive, and phonological resources available to the baby at the moment of production. First words often show subtle variation across different instances of production. For instance, the variants [baba], [βaβaba] and [baβa] may be produced for 'baby' or 'bottle' (Vihman, 2014; see also Thelen & Smith's (1994) discussion of the stability of motor routines involved in reaching and self-locomoting). This variation is an artefact of tensions between the baby's still-developing neuromuscular control, capacity for articulatory planning and cognitive abilities.

Vihman (1993) suggests that the articulatory filter may bias babies to select certain words to target from the input, guided by perceptual salience. In this way, the articulatory filter represents a powerful learning mechanism, giving babies a way to access parts of the speech stream and to engage with speech and language and behave more like those around them (Donald, 1991; Vihman, 2014). Previously, it was thought that babies explicitlyavoided producing words containing challenging or unfamiliar sounds. The articulatory filter offers an alternative account, which circumvents the psychologically unlikely idea that babies are able to render fully adult-like representations of target words and make decisions about which words to attempt or avoid according to the limits imposed by their own phonological systems (see Chomsky & Halle, 1968; Jusczyk & Aslin, 1995). Instead, the articulatory filter predicts that, once babies have enough productive, perceptual, and proprioceptive experience of producing certain sounds at will – that is, when they have a Vocal Motor Scheme – familiar sounds in the speech stream may capture their attention and they may be able to process these sounds top down as well as bottom up in real time (Vihman et al., 2014). Rather than simply processing the sensory perceptual information in the speech stream, babies may now be able to recognise and match up the sounds that other people make and the articulatory gestures that they themselves make to produce those sounds. Perceiving auditory information may activate proprioceptive sensory information about consonants that is stored in memory. Vihman (2002) has suggested that this may be underlain by activation of *mirror neurons*. These are neurons that are activated both when an individual (person or other primate) performs an action, *and* when they see the same activity being performed by another individual. Some research has found evidence of mirror neuron activity in humans when performing and observing motor actions, and some evidence of mirror neuron activity in processing the emotions of others (Fadiga et al., 1995; Bastiaansen et al., 2009; Rizzolatti & Sinigaglia, 2016).

An interesting association has been found between the number of Vocal Motor Schemes that a baby has and what parts of the speech stream are salient and interesting for them. Babies with a single preferred consonant prefer listening to speech that contains that sound, while babies with two or more preferred consonants show a shift in their attention towards speech that contains unfamiliar sounds (DePaolis et al., 2011; Majorano et al., 2014). This attentional shift is likely triggered by babies' increasingly well-developed ability to process familiar sounds in speech. As babies becomes more proficient at doing this, it requires less effort and attention, meaning that 'surplus' attention and energy is available for other new tasks. Once a baby has two or more stable consonants in repertoire, cognitive and sensory "processing resources" may become available, allowing them to learn to associate less familiar arbitrary phonetic signals (i.e., words) with their intended referents or meanings in the world (Ferguson & Farwell, 1975; Vihman & Croft, 2007; Vihman et al., 2009; DePaolis et al., 2011, p. 591; Vihman et al., 2014). This may trigger an adaptive reorganisation of a baby's developing phonological system, influenced by the baby's growing inclination to take part in social vocal interaction with those around them.

The articulatory filter, whilst not explicitly goal-directed, serves as a 'bootstrapping' mechanism by which babies can gain a foothold in speech and language acquisition (Vihman et al., 2009). Vocal Motor Schemes represent a control parameter, prompting babies to accumulate and match up perceptual and proprioceptive information about their own vocalisations and the ambient language. This, with time and experience, contributes to the emergence of whole-word segmentational ability. Vocal Motor Schemes motivate a dynamic, synergistic reorganisation of babies' vocal, rhythmic, motoric, and attentional capacities, which will gradually result in the emergence of a phonological system capable of acquiring adult language. While Vocal Motor Schemes seem to be endogenously motivated, babble does seem to be sensitive in some way to the ambient language. Recall that Arabic-acquiring babies have been found to produce laryngeals in their babble, while laryngeals seem to be 'phased out' for English-acquiring babies (Esling, 2012). The articulatory filter may also play a role in the perceptual narrowing that is necessary in order for babies to acquire the phonological inventory of their ambient language (Esling, 2012; Vihman, 2014). The articulatory filter may emerge out of an interaction between a biological predisposition for babies to behave like their conspecifics and experience-dependent interactions between the baby, their own body, and their environment (Donald, 1991; Mareschal et al., 2007).

Summary

Bates has described cognition and language as "a new machine that nature has constructed out of old parts" (Bates et al., 1979; Bates, 1999, p. 10; Bates, 2004, p. 250). By extension, the same may be true of babble. In this chapter, we have seen how babble evolves out of the vocal, motoric, and rhythmic behaviours that come before it and how it changes in form, becoming more sophisticated and well-controlled with time and articulatory experience. We have also seen how babble provides babies with powerful learning mechanisms that facilitate their eventual transition into word acquisition and how individual experiences with babble and related behaviours can shape each child's journey. General, non-linguistic capacities that we share with some non-human animals become bound together during babies’ phonological journeys resulting, eventually, in the emergence of two significant and uniquely human phenomena: canonical babble and language. Physiology that has evolved for purposes like self-feeding and protecting the airways acquires additional functions like vocalisation and vocal exploration. Cognitive capacities that evolved to ensure survival and cooperation in species with complex social structures like understanding turn-taking, attending to social vocalisation, and the drive to behaving like conspecifics allow babies to form associations between their own babble and other vocal behaviours and the speech of those around them. Through continuous reorganisation and ‘upcycling’ of capacities like phonation, rhythmic ability, and Vocal Motor Schemes – “old parts” already in stock in the dynamic history of their phonological system – babies come to implicitly “construct” new behaviours. There is a great deal more that can be said about how babble influences language acquisition and some of this will be discussed in Chapters 2 and 6 of this thesis. Chapters 2 and 3 will now introduce the debate surrounding the underlying nature of babble and introduce some neuroimaging and neurosurgical evidence relating to language, phonological development, laterality, and the brain.

Chapter 2 The underlying nature of babble

This chapter will discuss competing theoretical perspectives on what babble is and where it comes from. The Introduction chapter to this thesis considered what the study of babble might contribute to our understanding of the developmental and evolutionary origins of language. We will see, in this chapter, how bearing these goals in mind can help us to develop a theory of babble that is consistent with the principles of ontogeny and phylogeny. This chapter will be followed by a chapter discussing brain development in babies and the evolution of the human capacity for language.

"…language like every evolved form, is the product of successive ontogenies, its structure a record of its own evolution…"

(Studdert-Kennedy, 1990, p. 17)

Introduction

In the existing literature, there is some debate as to whether babble should be characterised as a 'linguistic' behaviour or a purely motoric behaviour (Lenneberg, 1967; Jakobson, 1968; Davis and MacNeilage, 1993, 1995; Petitto et al., 2004; Vihman et al., 2014). This debate has implications for a conceptualisation of language and language acquisition and these two competing accounts have somewhat complementary strengths and limitations. The *Motoric hypothesis* can account for the fact that babies' babble seems to be endogenously motivated (Oller et al., 1980, 2019). However, a purely motoric view of babble falls short of explaining how pre-existing vocal and motoric capacities become tuned towards language (i.e., the articulatory filter) and how their articulatory experiences and preferences (i.e., Vocal Motor Schemes) shape their transition into word form production (Menn, 1971;

Waterson, 1971; Priestly, 1977; McCune & Vihman, 1987; Keren-Portnoy et al., 2005; McGillion et al., 2017). These findings have been cited as motivation for the *Linguistic* hypothesis (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004). However, the Linguistic hypothesis is not without issue either. This view is associated with the maturational or Nativist view of language acquisition and situates language as a pre-determined endpoint for human phonological development, with stages in phonological development representing moves towards this endpoint (see Chomsky & Halle, 1968; Petitto et al., 2004). As discussed in Chapter 1, teleological and discrete stage-based theories of development are not well-supported by our current state of knowledge about psychology and development in biological organisms (Thelen & Smith, 1994; Mareschal et al., 2007). This view implies a degree of goal-directedness and intentionality, which babies under the age of one have not been observed to be capable (Thelen & Smith, 1993, 1994; Vihman, 2014). The Linguistic hypothesis does not offer a plausible explanation as to why babies babble or how babble emerges and changes over time.

The present thesis puts forward an original conceptualisation of the underlying nature of babble: the *Old Parts, New Machine hypothesis.* This hypothesis characterises babble, as described in Chapter 1, as a complex systematic phenomenon, emerging endogenously as a confluence of contributing behaviours and capacities, and which comes fortuitously to provide babies with phonological resources that pave their way into language, once the cognitive and attentional resources for 'doing' language become available. Unlike the Motoric hypothesis, this alternative hypothesis can account for how babies transition from babble into language. A key distinction between this view and the Linguistic hypothesis, is that language is not situated as a goal or trigger in phonological development (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004), but rather as the natural consequence of babies' growing phonological proficiency and cognitive development, along with their drive to behave like those around them (Donald, 1991).

This chapter will discuss these competing hypotheses in depth and consider not only how well they can account for phonological behaviour in typically developing babies and children, but also how well they can account for cases of atypical development and for the ontogeny and phylogeny of language. A satisfactory theory of the underlying nature of babble should be consistently applicable across these populations and levels.

The Motoric hypothesis

The Motoric hypothesis posits that babble is a purely motoric behaviour originating in much the same way as other motoric behaviours that emerge during the first year of life – that is, when motor control over the developing articulatory apparatus matures, babies may happen upon syllable production during vocal exploration (Lenneberg, 1967; Jakobson, 1968; Elbers, 1982; Kent, 1984; MacNeilage & Davis, 1993; Davis & MacNeilage, 1995; Locke, 2000; MacNeilage & Davis, 2000, 2001; Iverson & Fagan, 2004; MacNeilage, 2008; Fagan, 2009). As with other motor behaviours like kicking the legs or blowing 'raspberries', babies are suggested to find syllable production enjoyable or stimulating and may come to produce them again and again (Thelen, 1979, 1981). Another key factor in explaining babble from the perspective of the Motoric hypothesis is rhythmic development. Work discussed in Chapter 1 has explored the critical contribution of rhythmic ability to the adult-like timing of the consonant-vowel transition in syllables of canonical babble, and to the emergence of reduplicated and variegated babble sequences (Thelen, 1979, 1981; Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001; Iverson & Fagan, 2004; Iverson et al., 2007; Fagan, 2009). Some proponents of the Motoric hypothesis posit that, during babies' vocal exploration, their capacity for producing rhythmic, cyclical behaviours like kicking, rattle-shaking, and clapping becomes coupled with the capacity for vocalisation, and this gives rise to babble as a new form of vocal behaviour (e.g., Ejiri, 1998; Ejiri & Masataka, 2001). Indeed Ejiri and Masataka (2001) found that babble produced at the same time as rhythmic limb movements shows faster, more mature-sounding rates of production and consonant-vowel transitions similar to those observed in adult speech, whereas babble produced whilst the baby's body and limbs were still shows slower syllable production rates and less regular formant transitions. This is compelling evidence that rhythmic ability influences in-the-moment vocal production and, since everyday experiences shape and constrain development (Thelen & Smith, 1993), it is likely that rhythmic ability also exerts influence over a baby's wider phonological system over developmental time.

The Independence hypothesis

Some early proponents of the Motoric hypothesis suggested that babble is unrelated to later language acquisition (Lenneberg, 1967; Jakobson, 1968; Locke, 1983; MacNeilage &

Davis, 1993). Lenneberg (1967, p. 140) argues that vocal experience from cooing and babble does not constitute practice for language in any way, citing his own observations of one child who was tracheotomised between 8-14 months of age and who began to produce "babbling sounds typical of the age" no more than one day following decannulation. It is unclear from the context of Lenneberg's discussion what criteria he is using to define babble, and he does not comment on whether this child had begun to babble before their tracheostomy at 8 months. Neither does Lenneberg specify when this child began to produce words, nor whether there was any phonological similarity between their babble and their early word forms. Recall that babble typically emerges between 6-8 months – up to 2 months before this child underwent tracheotomy (Lenneberg, 1967) – and the first words typically emerge around 12 months of age (Oller, 1980; Vihman, 2014). Lenneberg (1967, p. 178) goes on to describe how phonological and linguistic milestone attainment is "interlocked" with the attainment of other motor milestones, suggesting that language acquisition, like motor skill development, is under maturational control. Jakobson (1968, p. 24) describes babble as "biologically oriented tongue delirium" and "the purposeless egocentric soliloquy of the child". He argues that babble has no bearing on or relevance for language acquisition and goes on to claim that a "silent period" may elapse between the time when babble is produced and the time when children acquire their first words, during which babies simply cease to vocalise. This latter idea was based on observations of a small sample of children and has not been well-supported by child language acquisition research. In fact, to the contrary, babble and word production have been found to co-occur – sometimes within a single utterance – for up to 8 months after first word emergence in typically developing babies,[8] and significant continuity has been observed between babies' idiosyncratic production preferences in their babble and their first words (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; Locke, 2000; Stoel-Gammon, 2001; Vihman et al., 2009; Vihman et al., 2014; Vihman, 2019).

---

[8] (and much longer in some atypically developing babies and children)

Alternative conceptualisations

Alternative conceptualisations of the Motoric hypothesis have recognised this continuity (Davis & MacNeilage, 1995; Locke, 2000) but have maintained that babble emergence is determined by maturation of the motor system and constitutes preparation for later language acquisition only in the sense that babble gives babies opportunities to develop stronger voluntary control over the articulators and to experience the auditory consequences – the sounds made – of the articulatory gestures that they produce (Elbers, 1982; Elbers & Wijnen, 1992; Elbers, 2000). Parallels can be drawn between the development of babble and that of rhythmic arm-waving, crawling, or walking (Thelen, 1979, 1981; Van der Stelt & Koopmans van-Beinum, 1986). Babble is not viewed as a goal-directed behaviour, nor is any notion of underlying innate linguistic neural mechanisms or activity invoked. Drawing on evolutionary principles, it is instead argued that babble arises from domain-general, non-linguistic capacities for movement, perception, and proprioception, and that the transition into language acquisition follows babble as an *a posteriori* consequence of the fortuitous availability of babble and other capacities (Studdert-Kennedy, 1990). Importantly, this alternative Motoric hypothesis does not seek to deny the critical role of language exposure and experience for *language* development. However, The Motoric hypothesis does not view *babble* as being 'triggered' by babies perceiving adult language (Elbers, 1982; Van der Stelt & Koopmans van-Beinum, 1986; Studdert-Kennedy, 1990; MacNeilage, 1998; MacNeilage & Davis, 2000; MacNeilage, 2008; Fagan, 2009). Nor does the Motoric hypothesis seek to deny the influence of perceptual and proprioceptive experience on the emergence of babble. Rather, this hypothesis views babble as one of the ways that babies respond to an endogenous, domain-general impulse to behave like those around them, which happens to afford them a certain amount of articulatory experience (Studdert-Kennedy, 1990; Donald, 1991).

Frame/Content theory

A highly influential theory under the Motoric hypothesis is MacNeilage and Davis' *Frame/Content theory* (MacNeilage & Davis, 1990, 1993; Davis & MacNeilage, 1995; MacNeilage, 1998). This theory proposes that syllables of babble – and adult spoken language – can be divided into two portions: the *frame*, or the cyclical movement involved in opening and closing the jaws, and the *content*, or the articulatory movement/s of the lips, tongue, and soft palate (Davis & MacNeilage, 1995; MacNeilage & Davis, 2000). These authors have found evidence suggesting that manipulation of the frame accounts for most of the phonetic variation in babble and word form production during the first 18 months of life, and that babies of this age have “extremely limited” neuromuscular control over their oral articulators (Davis & MacNeilage, 1995, p. 505; MacNeilage, 1998). The baby in MacNeilage and Davis' (1990) longitudinal single case study produced only around 60% of the vowels in her word forms faithfully to the adult target form, despite an exponential growth in her vocabulary, from 25 to 750 words between 14 and 20 months of age. The baby showed a strong preference for producing high front vowels adjacently to coronal consonants (e.g., [di]) and central vowels adjacently to labial consonants ([e.g., [bə]). The baby also showed a tendency to reduplicate both vowels and consonants when producing disyllabic words. Davis and Macneilage (1994) also observed these co-occurrences in earlier phases of phonological development, in the babble of one baby between the ages of 7 and 12 months. These babies were also observed to favour front sounds and produced back consonants and vowels much less frequently. These findings were replicated in another longitudinal study of six babies aged between 6-8 months and 3 years (Davis & MacNeilage, 1995). From the findings of these studies, Davis and MacNeilage concluded that the jaw is the primary articulator in babble regardless of whether babies are producing coronal or labial sounds. Davis and MacNeilage (1995, p. 1208) propose that reduplicated babble is produced by “uniform” repeated jaw oscillations, whilst variegated babble is produced by “non-uniform” oscillations of the jaw – uniformity here referring to the position of the articulators during the closed and/or open portion/s of a syllable.

Kinematic speech and babble research using OPTOTRAK technology has provided some support for Davis and MacNeilage’s (1995) claim that the mandible is the primary articulator in babble but not in speech (Munhall & Jones, 1998). OPTOTRAK is a specialised motion capture system that tracks movement of a person or object in 3-dimensional space in real time, using feedback from infra-red transmitters attached to the person or object (Munhall & Jones, 1988). Munhall and Jones (1998) attached transmitters to the jaws and lips of 8-month-olds and adults, and found that babies move their lips only as a consequence of jaw movement, while adults move their lips independently of their jaws during speech. Whilst these findings do lend support to Frame/Content theory, Munhall and Jones (1998, p. 154) explicitly state that they “restrict” their interpretation of their findings to observations about the present state of human language and babble. By contrast, Davis and MacNeilage have suggested that the *frame* then *content* pattern they describe in babble emergence and development may recapitulate the process by which human language evolved in our ancestors (MacNeilage & Davis, 2000; MacNeilage, 2008). A facial motion tracking study by Nip et al. (2009) found further evidence that articulatory control over the jaws and lips and tongue develops at different rates. Babies’ jaw oscillation speed was observed to stabilise and become adult-like by around 15 months of age, whilst neuromuscular control over the lips remained immature, beginning to stabilise only around 18 months of age.

Returning to the phylogeny of babble, under Frame/Content theory, the frame is theorised to originate from the repeated cycles of jaw opening and closing during nutritive and non-nutritive sucking (MacNeilage, 1998). MacNeilage (1998, p. 504) notes that members of other species like baboons and chimpanzees seem to combine feeding-associated oral gestures like lip-smacking and tongue-flicking with phonation for social communication. MacNeilage proposes that this capacity for combination and modification and our tendency to exploit it was inherited from our shared ancestor with these non-human primates, and he suggests that babies may make similar use of pre-existing oral gestures from feeding during vocal play. He suggests that, during vocal play, some of these gestures may be involved in babies’ ‘discovery’ of babble (MacNeilage, 1998, 2008). This idea parallels some of the findings relating to vocal learning, contingent feedback, and call convergence in non-human primates and other animals.

Frame/Content theory and language evolution

MacNeilage (2008) and other authors have further argued that ingestive gestures may have been instrumental in the evolution of vocal communication and language in humans and that, in this way, Frame/Content theory may also account for the adult state of human language as a product of the type of vocalisations that come naturally to babies. MacNeilage (2008) suggests that these vocal ingestive gestures may first have become associated with meanings when they were produced during caregiver-infant dyadic interactions either by one interlocutor or the other, and further suggests that such vocalisations had benefits as a vehicle for social bonding, and for ensuring social co-operation and survival. This argument is consistent with ideas about self-constraining and self-determining development in biological systems. However, this theoretical strength does not wholly immunise the Motoric hypothesis and Frame/Content theory against criticism.

It has been suggested that syllables in many world languages adhere to some of the motoric constraints described by Frame/Content theory (Locke, 1983; Studdert-Kennedy, 1990; Locke & Pearson, 1992; Davis & MacNeilage, 1995; MacNeilage, 1998). For instance, in babble, closures at the front of the oral tract are common and are often produced in the phonetic environment of a low front vowel. It is argued that when front consonants and low front vowels co-occur, the tongue remains in or near resting positions, with its vertical movement within the mouth being modified primarily by jaw movement (Studdert-Kennedy, 1990; Davis & MacNeilage, 1995). Common production preferences like this apparent constraint on tongue movement can be seen in many world languages: apical stops usually co-occur with front vowels and dorsal consonants with back vowels (Locke, 1983; MacNeilage & Davis, 2000). Furthermore, a subset of consonants containing /p, b, m, t, d, n, k, g, h, w, j/ represents 90% of the consonants produced in that babble of 12-month-olds, and these same consonants are found in the phonological inventories of the majority of world languages – even those with relatively small consonant inventories (Lindblom & Maddieson, 1988, cited by Locke & Pearson, 1992; MacNeilage, 1998, p. 505). Indeed, looking again at Table__, we see that these consonants were also those most frequently settled on as Vocal Motor Schemes by babies in the present study. By contrast, palatals, liquids, affricates and fricatives are less common in babble and in world languages (Locke, 1983; Locke & Pearson, 1992). Indeed, stops, nasals and glides appear in 90% of world languages, whilst oral fricatives appear in only 50% and affricates in only 40% (Locke & Pearson, 1992)

MacNeilage and Davis (2000, p. 529) argue that these tendences reflect “fundamental” biological constraints, existing early in language evolution and that ease of articulation may have resulted in more frequent production of certain sounds, increasing the likelihood that these sounds would become associated with frequently encountered or required meanings and referents. Considering /p, b/ as an example, as well as being relatively easy for babies to produce, the jaw movement that creates the frame for these sounds may represent a more visually and proprioceptively accessible target compared with sounds that involve making contact in less visible parts of the vocal tract such as the hard palate (McCune, 1998; Keren-Portnoy et al., 2005; DePaolis et al., 2013). This frame can be extended to allow a speaker or babbler to produce /t, d/ by slightly altering the resting position of the tongue in the mouth. This fact may account for the prevalence of both /p, b/ and /t, d/ in world languages. In language evolution, jaw oscillation may have been a valuable resource for mutual interaction between caregivers and babies, and between adults. Through repeated production in feeding and in caregiver-infant interaction, the jaw motion involved in producing these sounds may become well-practised and quite automatic. Automaticity can free up cognitive and motor planning resources, making these resources available for other functions like conveying and interpreting social information and processing meaning and reference.

MacNeilage (1998) argues that the fact that wild and captive chimpanzees are observed to produce voiced lip-smacking gestures (which involve similar open-close movements of the jaw to labials and apicals) during social grooming episodes lends further support to this idea. In chimpanzees, production of these lip-smacks during social grooming has been found to be associated with greater co-operation and longer episodes of social grooming (Fedurek et al., 2015a). MacNeilage and Davis (2001, p. 79) suggest that gestures like lip-smacking may represent an “intermediate stage” in the evolution and development of the protosyllable for communicative and linguistic vocalisation: they believe that gestures like lip-smacking may have provided the frame, to which could then be added voicing and supraglottal content. Interestingly, MacNeilage (2008) further notes that chimpanzees show not only frame production and voicing, but also evidence of conceptual understanding, meaning that several of the ‘raw ingredients’ for language were likely available to our common ancestors with chimpanzees. However, chimpanzees apparently do not show any evidence of producing consonant- and vowel-like content within these frames (MacNeilage & Davis, 2001).

MacNeilage and Davis (2001) go on to cite work by Donald (1991) on mimesis and work by Meier et al. (1997) on silent jaw-wagging in support of the foundational claim of Frame/Content theory: that babble emerges from development in general non-linguistic capacities. Meier et al. (1997) found that some 8-13-month-olds, both hearing and hearing impaired, can be observed to produce rhythmic cycles of jaw opening and closure without voicing or breathing – a silent frame in the terms of Frame/Content theory. These jaw wags were produced sometimes in isolation and sometimes in the context of a sequences of oral gestures which also contained some voiced canonical babble. MacNeilage and Davis (2001) take this finding, alongside MacNeilage et al.'s (2000) observation that babble is characterised by CV syllable structure and is often reduplicated unlike most words in most world languages, as evidence that babble originates endogenously rather than being influenced by caregivers and others modelling adult language.

Donald (1991) has suggested that aspects of human cultural behaviour like language, paralinguistic cues, and dance are transmitted through mimesis. Mimetic ability is a general capacity that permits individuals to understand what is represented by a behaviour that they observe being produced, to create an internal representation associated with that behaviour, and then to reproduce that behaviour spontaneously and intentionally or invent related behaviours. Mimesis is related to but critically distinct from mimicry and imitation. Mimicry involves only precise copying, whilst imitations may involve spontaneous production but do *not* require representational ability (Donald, 1991). Donald explains how mimetic ability likely emerged in our ancestor *Homo erectus* and is heavily implicated in social co-operation, pedagogical interaction, and the building of culture in our evolution. MacNeilage and Davis (2001) interpret Donald's work as support for the idea that babble emerges through the coupling of a baby's endogenous drive to vocalise rhythmically, and the drive to behave like those around them. The argument is that babble is underlain by this general mimetic capacity that is sensitive to the domain-general behaviour of others, rather than some capacity that is specifically sensitive to language and its properties (Thelen, 1981; MacNeilage & Davis, 2001). Indeed Donald (1991) notes that whilst mimesis is a multimodal capacity, most of the behaviour that is transmitted from one human to another through mimesis – like, for example, facial expressions or tool use – is transmitted through the visual and motor/haptic channels. Furthermore, Donald (1991) states that some capacities traditionally implicated in language, like intentional communication or processing referential meaning are also inherently involved in mimesis, and suggests that the capacity for speech and language may draw on these pre-existing general, non-linguistic, mimetic capacities. Donald (1991, p. 165) describes how evolution – and development – "conserves" functionally beneficial adaptations from previous stages and argues that the mimetic ability necessary for *Homo erectus* to develop their complex technology and their social and pedagogical cultures should have left some “cognitive vestiges” in their descendants – including us – since this capacity can be used to transmit and acquire skills and behaviours that humans make use of day to day. This evolutionary principle is echoed by Studdert-Kennedy (1990, p. 17) in his discussion of the ontogeny and phylogeny of language: “language like every evolved form, is the product of successive ontogenies, its structure a record of its own evolution”.

So, a fundamental proposition of Frame/Content theory and the Motoric hypothesis more widely is that babble emerges endogenously rather than as a response to language exposure. Proponents of this hypothesis argue that a theory of the underlying nature of babble can and should be consistent with principles of evolution (MacNeilage, 2008). Considered alongside the evidence regarding the emergence of babble discussed in Chapter 1, this hypothesis seems to have some considerable strengths. The following sections will now explore some weaknesses and shortcomings of the Motoric Hypothesis and Frame/Content theory.

Criticisms of Frame/Content theory

Whilst it offers a conceptualisation of babble consistent with current theories of evolution and infant development, the Motoric hypothesis has faced some criticism from other researchers in phonological development and language acquisition. Some of these criticisms have addressed finer points of Frame/Content theory (e.g., Andrew, 1998; Bloom 1998; Ghazanafar & Katz, 1998; Jürgens, 1998; Lund, 1998; Menn, 1998; Ohala, 1998), and these will be discussed in the next sections. In particular, MacNeilage, Davis, and colleagues have made some claims that are not consistent with some of the research on babble development. Other criticisms have contested some of the more foundational assumptions of the Motoric hypothesis. These latter questions motivated the formulation of the Linguistic hypothesis and will therefore be discussed in the following section.

Well-formedness

In particular, MacNeilage and Davis (2001, p. 79) have stated that the frame of babble "emerges more or less fully formed rather than being put together in the uncoordinated manner characteristic of many baby actions sequences (e.g., using culinary utensils)". They have cited the fact that caregivers can reliably identify when babble emerges (Koopmans-Van Beinum et al., 2001; Oller et al., 2001) as support for this claim. However, Oller (1980) describes how, before the emergence of canonical babble, babies can be observed to produce marginal babble at around 4-6 months of age, wherein consonant-like closures are paired with vowel-like articulations with adult-like resonance. Critically, in marginal babble, the timing of the transition between the consonant- and vowel-like elements is *not* adult-like. This transition is slower, and its phonetic quality reflects babies' as yet immature articulatory planning and neuromuscular control (Kent, 1993). That is to say, this transitional behaviour *is uncoordinated* compared to consonant-vowel transitions in adult speech. The findings of Fagan and Iverson's (2007) mouthing study may also give further cause to question this claim about instantaneous well-formedness. The authors found evidence that 6-9-month-olds produced more supraglottal consonants when babbling alongside mouthing their hands and other objects and argue that mouthing may help babies learn about forming supraglottal closures. Mouth-hand or mouth-object closures are critically different from intentionally produced consonant-vowel vocalisations in terms of a) the baby's goal in producing the action and b) the movements of the jaw involved in producing the movement. Mouthing satisfies a drive for sensory exploration and does not usually involve creating a complete closure within the vocal tract itself unlike consonant production, chewing, or sucking (Fagan & Iverson, 2007). It is therefore not necessarily accurate to claim that syllables emerge "fully formed" (MacNeilage & Davis, 2001, p. 79) though it may appear to happen in isolated cases where babies' babble emerges late or their coordination stabilises rapidly or early. It may be more accurate instead to say that what caregivers can readily identify is the point when babies have accumulated enough articulatory practice to transition from the relatively 'sloppy' syllables of marginal babble to producing more adult-like canonical syllables. The cyclical motion involved in nutritive and non-nutritive sucking, which is established during the early days and weeks of life, may in fact not provide babies with enough articulatory experience to begin to produce the "fully formed" consonant-vowel alternations of canonical babble, since these emerge only after a period of producing the less ‘well-formed’, more uncoordinated consonant-vowel alternations of marginal babble (Bloom, 1998; Menn, 1998). Additionally, MacNeilage and Davis (2001) have claimed that babies do not produce consonants with no accompanying vowel, whilst Oller (1980) and Elbers (1982) have claimed that babies may be observed to produce isolated consonants from around 0-1 months of age.

The origins of the jaw motion

It has been suggested by Andrew (1998) that the cyclical jaw movements of babble may originate not in feeding behaviour, but in vocal behaviour. Andrew argues that the ability to repeatedly open and close part of the vocal tract (like the jaw) during a voiced breath sequence might easily follow when an animal with the capacities to plan action and to perceive and act upon feedback produces a single open-close alternation and then plans to repeat this. By the time that babble emerges, babies already show some developing rhythmic ability in other domains like arm-flapping and some growing ability to produce oral gestures intentionally, though some vocalisations still appear to be incidental rather than planned. It could be argued that rhythmic jaw movement involved in eating may still exert some influence over babble emergence, even if it cannot be said to be wholly responsible for it. It could also be argued that the pre-existence of rhythmicity in the jaw movement for eating might at least contribute to or influence the form of babies’ jaw movements during babble since behaviours ‘stored’ in the dynamic history necessarily influence the way that other behaviours using the same parts of the body are produced (Thelen & Smith, 1994; Lindblom, 1998). Lund and Kolta (2006) describe how a common Central Pattern Generator in the brainstem may generate the rhythm of both chewing and speech, arguing that the existence of clusters of neurons whose function is to rhythmically innervate the tongue, lips, and teeth during feeding leaves little natural motivation for a separate Central Pattern Generator to form for speech.

Ohala (1998) observes that chewing involves lateral motion that is not observed in speech and that syllables in speech often involve little or no jaw opening, in contrast to the jaw opening seen in chewing (see also Jürgens, 1998). However, in typically developing 6-9-month-olds and some premature babies who have received non-nutritive sucking interventions to support normal feeding development, nutritive and non-nutritive sucking has been found to show more *vertical* than lateral motion of the jaw and tongue (Miller & Kang, 2007). So, whilst the rhythmic jaw movement of chewing may be of questionable validity as a proposed source for the jaw movements in babble or speech, the rhythmic jaw movement of sucking could represent a more likely candidate, since the motor routine for sucking is already in repertoire before babble emerges at 6-8 months (Oller, 1980). However, kinematic research has found electromyographic evidence that babies aged between 8 and 22 months show distinct and unrelated patterns of coordination during sucking, chewing, and babbling (Moore & Ruark, 1996; Steeve et al., 2008; Steeve, 2010), which the authors believe result from differences in demands posed by each of the behaviours. Significant differences were observed between sucking and chewing, but also between sucking and babbling. Furthermore, muscle coordination and activation during all three behaviours was found to be more variable in younger babies than older babies or adults, suggesting that even at 8 months, babies' neuromuscular control over chewing is not yet "fully formed" (MacNeilage and Davis, 2001, p. 79) or adult-like. The findings of these studies do not support theories positing that babble and speech emerge out of nutritive jaw movements.

Another counter-argument against Frame/Content theory and the Motoric hypothesis more generally from Ohala (1998) may be that, unlike chewing, babble and speech show influences of auditory feedback. Similarly, Menn (1998) argues against the ideas that jaw movement – the frame – is the primary source of phonetic variability in babble, pointing out that auditory exposure may give rise to differences in the vocal behaviour of hearing and hearing-impaired babies. Menn (1998, p. 534) cites research by Wallace and Yoshinaga-Itano (1997), finding that hearing babies tend to favour oral stops in their babble, whilst hearing-impaired babies produce "almost *no* jaw-based rhythmic babble" but any babble that they do produce tends to contain only glottal stops or fricatives. These sounds provide hearing-impaired babies with an alternative source of feedback – haptic and proprioceptive feedback. Jaw, lip, and tongue motions from earlier vocal behaviours like *goo* sounds and vocalic shouts (Oller, 1980) may provide hearing babies with more directly relevant articulatory experience that is more likely to influence the shape babble. The role of auditory and proprioceptive feedback in babies' vocal behaviour will be explored in more depth later in this chapter.

Frame/Content theory, babble, and the brain

Several authors have agreed with the general approach of Frame/Content theory in that it draws on evolutionary principles and posits that speech and language evolved from general, non-linguistic capacities, but have argued that MacNeilage's (1998) claims about the neural activity supporting frame and content production is oversimplified or inaccurate (Abbs & DePaul, 1998; Ghazanfar & Katz, 1998; Lund, 1998; Sessle, 1998). Ghazanfar and Katz (1998) agree with the assertion that the frame and content portions of syllables can be meaningfully divided. Lund (1998), by contrast, argues that, if jaw, lip, and tongue movement during ingestion are controlled by the same neural architecture, the position that jaw movement should be supported by wholly different brain region/s to lip and tongue movement during babble or speech is unjustified. However, where Lund (1998) and Ghazanafar and Katz (1998) agree with one another is in that the assertion that syllable frames are produced by the medial premotor system and content is produced by the lateral premotor system, which is involved in chewing in non-human primates is problematic (Abbs & DePaul, 1998; Ghazanfar & Katz, 1998; Lund, 1998; Sessle, 1998). Counter to MacNeilage's (1998) claim that non-human mammal vocalisations do not contain rhythmic cycles, Lund (1998) states that the medial cortex is also involved in vocalisation in non-human animals. This same part of the brain, MacNeilage (1998) claims, is involved in rhythmic frame production in humans. Sessle (1998, p. 529) notes that there is pre-existing evidence that the lateral frontal cortex is, in fact, involved in "initiat[ing] and control[ling]" movements of the supraglottal vocal tract including the jaw. This part of the brain, MacNeilage (1998) claims, is responsible only for the content portion of syllables. Abbs and DePaul (1998, p. 511) go further in stating that "MacNeilage has lumped together the functional characteristics across multiple mesial and lateral motor cortex fields, inadvertently creating two hypothetical centres that simply may not exist". Abbs and DePaul (1998) also state that lateral rather than medial areas of the brain are involved in both ingestion and rhythmic coordination of speech, and that it is the lateral pre-central cortex rather than the medial pre-central cortex that contains projections to the parietal language areas.

These authors (Abbs & DePaul, 1998; Ghazanfar and Katz, 1998; Lund, 1998; Sessle, 1998) also concur amongst themselves (in opposition to Davis and MacNeilage) that movement of the jaw and other articulatory apparatus is underlain not by activation in one or two particular area/s of the brain but instead by activity distributed across a network of several brain areas. Ghazanfar and Katz (1998, p. 516) make the point that the lateral and medial brain areas in primates have “extensive overlapping and reciprocal connections with other cortical areas involved in orofacial movements (Luppino et al., 1982; Matelli et al., 1986)”. Sessle (1998) notes that electrical stimulation of four different areas in the brain areas gives rise to four slightly different rhythmic motions of the same part of the body – namely, the jaw. The areas in question are: the face area of the primary motor cortex; the face area of the somatosensory cortex; the cortical masticatory area that sits alongside the primary motor cortex; and a second cortical masticatory area that sits below the frontal operculum. Ghazanfar and Katz (1998) suggest that the Central Pattern Generator responsible for rhythmic jaw movement may be sufficiently general to produce both chewing and babble/speech precisely *because* it relies on such a distributed network of brain areas.

Frame/Content theory and frequency

Jürgens (1998) has argued that the difference in frequency between speech syllable production (around 5Hz) and chewing (around 1.5 Hz) casts doubt on ingestion as the source of the rhythmic cyclicities of adult speech. Research by Zimmerman and Foran (2017) found that non-nutritive sucking has a frequency of (2Hz) similar to that of chewing. It is possible that this discrepancy in frequency may be explained in terms of the mechanics of feeding compared with vocalising. Feeding involved moving matter from the mouth into the stomach, so the cyclical jaw movements for ingestion must be coordinated with swallowing and breathing. Additionally, these jaw movements are met with resistance from the matter being eaten, which may slow down the rate of jaw oscillation. Differences in the texture of a substance being ingested are associated with differences in jaw motion and degree of neural activation (Onozuka et al., 2002). Babble and speech involve neither coordination with swallowing nor this type of resistance. Decreased physical effortfulness may mean that energy is therefore available to produce jaw movements at a faster rate, allowing the speaker or babbler to produce more sounds before the ‘fuel’ for voicing (the breath) runs out. If this is the case, differences in the frequency of speech/babble and ingestion may not pose a significant problem for Frame/Content theory. Nip et al. (2009) found evidence that babies show an adult-like rate of jaw movement in babble and first words but a slower rate of their silent spontaneous orofacial behaviours other than feeding, like silent jaw wags. However, research reviewed by Ghazanfar and Takahashi (2014) has found that certain vocalisations of rhesus macaques are produced with a rhythm between 3-8Hz – much closer to that of speech than of feeding behaviours. The authors suggest that the rhythm of spoken language is more likely to have emerged from the rhythm of pre-existing vocalisations and facial expressions in our development and our evolution than from feeding behaviours.

Summary

To summarise, the Motoric hypothesis posits that babble is underlyingly a purely motoric behaviour whose development is comparable to that of other motoric behaviours. It is argued that babble emerges through endogenously driven vocal exploration and its form is determined and constrained by the range of motion that is motorically available to babies at any given time in their physiological and neuromuscular development. Similarly, some proponents of the Motoric hypothesis have also suggested that tendencies in adult language arise out such motoric constraints (). The Motoric hypothesis offers a view of babble that is somewhat plausible in evolutionary terms, in that it describes a bottom-up process of development and perhaps also of evolution, and it provides some possible answers as to how language may arise out of pre-existing general capacities. However, discrepancies have been noted between the claims made by two key proponents of the Motoric hypothesis – Davis and MacNeilage (1993, 1995, 1998, 2008) – and the findings of research relating to rhythmic jaw movement, phonological and neuromuscular development, and neuroanatomy. Additionally, one area where the Motoric hypothesis may be lacking is in its ability to account for the *top-down* processes that can be observed to take place when babies make the transition from babble into first words. Critics of the Motoric hypothesis and proponents of the Linguistic hypothesis have demonstrated how the Motoric hypothesis fails to account for how and why ambient language exposure can shape the phonological form of babble; how and why babies develop at markedly different rates; and how and why babies develop individual production preferences (Elbers, 1982, 2000; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; de Boysson-Bardies & Vihman, 1991; Andrew, 1998; Bloom 1998; Ghazanafar & Katz, 1998; Jürgens, 1998; Lund, 1998; Menn, 1998; Ohala, 1998; DePaolis et al., 2011, 2013; Majorano et al., 2014; Vihman 2014). Since some of these criticisms are foundational to the formulation of the Linguistic hypothesis, they will be discussed in the following section

The Linguistic hypothesis

The Linguistic hypothesis is aligned with the Nativist view of language acquisition and characterises babble as a “fundamentally linguistic” mechanism for facilitating language acquisition (Petitto & Marentette, 1991; Holowka & Petitto, 2002a, p. 44; Petitto et al., 2004). This view argues that babble results from the maturation of some “brain-based language capacity” and is “triggered” when babies perceive the phonetic units, syllables, and rhythm of language (Petitto & Marentette, 1991, p. 1495; Petitto et al., 2004, p. 46). This “brain-based” capacity is believed to make humans babies uniquely sensitive to these properties of language and this allows babies to segment the speech stream and learn to produce speech-like syllables. According to this hypothesis, babble emerges when babies perceive the syllables of adult language and are motivated to try to produce speech-like syllables themselves (Holowka & Petitto, 2002a; Petitto et al., 2004).

The Linguistic hypothesis has been developed in response to the Motoric hypothesis (Petitto et al., 2004). As mentioned earlier, early conceptualisations of the Motoric hypothesis that claimed that there was no relationship between babble and early language acquisition are not well-supported by research finding continuity in the phonological properties of babble and first words (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; de Boysson-Bardies and colleagues 1981, 1989, 1991, 1993; Elbers and colleagues 1982, 1985, 1987, 2000; Locke, 2000; Keren-Portnoy et al., 2005; Vihman and colleagues, 1985, 1987, 2009, 2014, 2019; DePaolis et al., 2011, 2013). Similarly, the Motoric hypothesis’ claim that babble is not influenced by properties of the ambient language is also not wholly supported (DePaolis et al., 2011, 2013; Majorano et al., 2014). If babble were entirely determined and constrained by the maturation of the motor system, we might expect a) that consonants should emerge in the same order in all babies, and b) that all babies should settle on the same, motorically-determined favoured consonants or *Vocal Motor Schemes* (see McCune & Vihman, 1987). However, as discussed in Chapter 1, this is not what we observe. The

Linguistic hypothesis was proposed to account for how babies develop idiosyncratic phonological tendencies and preferences; for how these preferences shape their journey into language acquisition; and for the ways in which social interaction, language, and contingent feedback can be observed to influence babies' babble (see discussion in Petitto et al., 2004). The following pages will review some of the research evidence that has been cited in support of the Linguistic hypothesis.

Triggering and systematicity

A central claim of the Linguistic view is that, rather than emerging endogenously, babble is instead 'triggered' when babies perceive the patterns, contrasts, and rhythms of language and try to produce these themselves (Petitto & Marentette, 1991; Petitto et al., 2004). Proponents argue that, whilst babble does involve a motoric component, babies' apparent sensitivity to these properties of language and the ways in which this sensitivity is perceived to determine babble constitute evidence that babble is "systematic and fundamentally linguistic" (Petitto et al., 2004, p. 47).

This claim may be supported by research finding that babble, like language, may be modality-independent. Both deaf and hearing babies who are systematically exposed to sign language may produce some hand gestures which have been termed *manual babble* or *sign babble* (Petitto & Marentette, 1991; Petitto et al., 2004). Petitto and colleagues (1991, p. 1495; 2004, p. 69) state that sign babble is "identical" to vocal babble: like vocal babble, it is produced with adult-like rhythm (reflecting that of adult sign language), without meaning or reference and is reduplicated, citing Oller and Eilers' (1988) criteria for identifying babble[9] (Petitto & Marentette, 1991; Petitto et al., 2004, p. 1495). Additionally, sign babble is reportedly produced within the *sign space* – the area in front of the torso and head (Petitto & Marentette, 1991; Petitto et al., 2004, p. 1495).

---

[9] The discrepancy between Petitto and colleagues' (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004) reported criteria for defining babble and Oller and colleagues' (Oller, 1980; Oller & Eilers, 1988; Oller, 2000) criteria has been discussed in Chapter 1 and will be revisited later in this chapter.

In a study involving two sign exposed deaf babies of deaf parents and three speech-exposed hearing babies of hearing parents, recorded at 10, 12, and 14 months of age, Petitto and Marentette (1991) found evidence to suggest that the manual movements of sign-exposed babies may show some systematicity. Between 32-71% of the sign-exposed babies' manual movements were classified as manual babbles, whilst this proportion was lower for speech exposed infants, at 4-15%. Of the manual movements classified as sign babble, these contained 13/40 of the hand shapes used in ASL and 98% were produced in the sign space. However, only 47% of these movements classified as sign babble met the authors' stated criteria for identifying sign babble – namely, reduplication. There appears to be some inconsistency or contradiction here. However, recall that reduplication is *not* in fact a defining characteristic of babble (Oller, 1980). So, amongst the remaining 53% of hand movements that Petitto and Marentette labelled manual babble despite not meeting all of their criteria for manual babble, there may have been some hand movements that would be classifiable as babble under Oller's (1980; 1988; 2000) original full criteria. Petitto and Marentette (1991) took their findings to be evidence that exposure to sign gives rise to more systematic manual behaviour – sign babble – and, by extension, that sign exposure constitutes language input that motivates babies to babble using their hands rather than their voices. If true, this would lend support to the idea that linguistic input triggers babble emergence and production.

However, the fact that hearing babies not exposed to sign language produce hand movements that meet the criteria for manual babble gives cause to question the argument that manual babble must be triggered by sign exposure. In Petitto and Marentette's (1991) study, these manual babble hand movements made up 4-15% of hearing, non-sign-exposed babies' hand movements. Petitto and Marentette (1991) attribute this to to accidental production. The authors liken these hand movements to the vocal babble produced by deaf babies in that they are contain a more limited set of hand configurations than the sign babble of sign-exposed babies (c.f. Oller & Eilers, 1988; Petitto & Marentette, 1991; Petitto et al., 2004). Petitto et al. (2004, p. 46) argue that the suggestion by proponents of the Motoric hypothesis that *vocal* babble may first emerge endogenously as a "natural accident" or a "biological "side effect"" is psychologically unlikely. Petitto and colleagues do not, however, acknowledge any similar problem or inconsistency with their own earlier suggestion (see Petitto and Marentette, 1991) that hearing impaired babies happen upon vocal babble and hearing babies exposed only to speech happen upon manual babble purely accidentally. Additionally, in the latter of these papers, Petitto et al. (2004, pp. 47, 69) make the following claims, disregarding the 4-15% occurrence of sign babble in speech-exposed hearing infants:

“In the course of examining profoundly deaf babies exposed to a sign language, Petitto and Marentette (1991) observed a class of hand activity that was like no other. It was not like the deaf babies’ gestures nor anything else that they did with their hands; nor was it like any class of hand activity observed in the hearing control babies.”

“A final puzzle is this: do hearing babies acquiring spoken language produce manual babbling as seen in babies acquiring sign languages? No.”

Whilst it is clear that rhythmic ability plays a critical role in rhythmic hand movement, Petitto and Marentette (1991) do not discuss general rhythmic development. The authors also do not comment on whether any babies in this sample had begun to produce meaningful signs. This could be informative since babble and word production have been found to exert influence over one another and having some meaningful signs in repertoire may affect the way that sign-exposed babies produce sign babble (Elbers, 1982; Elbers & Ton, 1985).

In their later study, Petitto et al. (2004) used OPTOTRAK software to examine the frequency of manual movements (i.e. the number of cyclical hand movements per second) made by six hearing babies, three of whom were exposed to speech and three to sign language, at ages 6, 10, and 12 months. The study found that both speech and sign exposed babies produced general manual movements at around 2.5-3Hz (i.e., 2.5-3 cyclical hand movements per second) whilst sign exposed babies produced manual movements at around 1Hz more frequently than did speech exposed babies. Sign exposed babies showed a general but not absolute tendency to produce manual movement meeting some or all criteria for sign babble at lower frequencies (Hz) than other manual movements. This was taken by the authors to indicate that sign exposed babies were exhibiting two underlyingly different types of manual behaviour. Whilst speech exposed babies did produce some low frequency manual movements, the authors report that none of these were coded as manual babble and claim that no more than 8% were produced in the sign space. This seems surprising, since the sign space represents a fairly large proportion of a baby's possible range of motion. This fact that manual babble and non-babble manual movements showed differences in frequency is taken by the authors as evidence that sign babble is linguistic rather than motoric. Petitto et al., (2004, p.70) claim that this "could only" arise if babies possess some brain-based language processing mechanism that is triggered when babies perceive language (sign, in this case), and that motivates them to copy language and equips them with the motoric resources to do so (Petitto et al., 2004, p. 70).

However, a paper titled *How Fast is Sign Language? A Re-evaluation of the Kinematic Bandwidth Using Motion Capture* found evidence that the majority of sign language may occupy *higher* rather than lower frequencies (Bigand et al., 2021). The authors note that, whilst previous studies examining human arm and hand movement have estimated that isolated signs and finger spelling used in sign languages occupy a bandwidth of 0-5Hz, 0-7Hz, or even 0-3Hz (Sperling et al., 1985, Poizner et al., 1986, and Foulds et al.,2004, respectively, cited in Bigand et al., 2021), this bandwidth fails to accommodate the full range of frequency of motion found in fluent signed language. Rather, Bigand et al. (2021) found that continuous, contextualised sign language occupies a much wider bandwidth of 0-12Hz. It is notable that even this outdated range of 0-3Hz would encompass both the "linguistic" and "non-linguistic" manual movements that Petitto et al. (2004, p. 43) describe. Were babble were linguistic and were it the result of the triggering of an innate linguistic brain mechanism, then we might expect to see the reverse pattern to that seen in the manual babbles of the babies in Petitto et al. (2004) study, with manual babble being produced at a *faster* rate than other non-babble manual movements.

The argument regarding triggering is that if the capacity and tendency for babies to babble is modality independent and is triggered by perceiving language, then this suggests that it may rely on activity in part or parts of the brain that are evolved to process language. Petitto and Marentette (1991, p. 1493) conclude that "babbling is tied to an abstract linguistic structure of language and to an expressive capacity". However, there is another way to interpret these findings: human babies, much like the infants of non-human animals, may be biased to behave like their conspecifics and this bias may be expressed across multiple modalities (Donald, 1991). It may simply be that babies begin to produce certain motor actions during exploratory play and only later, come to recognise similarities between what they themselves do with their own hands or mouth and what their caregivers do with theirs. Babies may then be motivated to repeat particular motor actions, not by a drive to acquire or use language, but simply by a domain-general drive to do what those around them do. It may be that patterned linguistic input is sufficient but not necessary to motivate babies to do this.

Petitto et al. (2004, p. 49) draw a distinction between manual babble and other motor and rhythmic movements, arguing that the emergence of rhythmic stereotypies at 5-6 months, and their decline at 9-10 months, along with the fact that they seem to require no "clues from the input", indicates that they are under maturational control. As we have seen in Chapter 1, it seems that some sensory information from the input is important in babble. However, it is not necessary to posit that this sensory information is involved in triggering babble. Certainly sensory information from both the input and the baby's own output (Elbers, 2000) seems to be important in motivating babies to *repeat* certain behaviours, including babble. However, no comprehensive explanation is given as to why babies must perceive language in order to *begin* to babble. As Petitto and colleagues have noted, deaf and sign exposed babies may happen upon canonical syllables having never heard speech, and hearing, speech-exposed babies may happen upon sign babble having never seen signed language. In cases like these, something is 'triggering' or causing the behaviour at the very first instance of production, and it cannot be language exposure.

Continuity between babble and words

Continuity between babble and speech is well-documented (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; Elbers, 1982; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; Elbers, 2000; Locke, 2000; Keren-Portnoy et al., 2005; Vihman et al., 2009; DePaolis et al., 2011; DePaolis et al., 2013; Vihman et al., 2014). Evidence from studies like these has been used in arguments against the Independence hypothesis, since they demonstrate that what babies do in babble can significantly influence what they later do in language.

Oller et al. (1976) identified significant similarities in babies' production preferences in their babble and early word forms, both in terms of sounds and apparent phonotactic constraints. They found that 90% of consonant positions in babble were occupied by singleton consonants, and this was reflected in their early word form production. These early forms included target forms containing singleton consonants as well as target forms containing consonant clusters which babies reduced to singleton consonants. Additionally, Oller et al. (1976) observed that babies produced syllable-initial stops 10 times more frequently than syllable-initial fricatives and affricates, and produced fricatives and affricates 3 times more frequently than stops in syllable-final position. These position preferences were reflected in the early word forms of babies in this study. Vihman et al. (1985) also examined consonant distribution in babble and first words during infant-caregiver dyadic interactions and during solitary play in a longitudinal study involving nine babies aged 9-16 months. They observed some parallel development in the phonotactic structure of babies' babble and first words, with single vocalic (V) productions being following by consonant-vowel (CV) syllables, and later by (consonant-)vowel-consonant ((C)VC) syllables. The authors also found evidence that consonants emerging in babble often spread into speech. For instance, relatively few babies in their sample (3 out of 9) frequently produced velar consonants in their babble, but those who did later frequently produce words containing velar sounds.

The observation that consonants produced in babble also appear in first words lends some support to the Linguistic hypothesus. Vihman et al.'s (1985) finding that babies frequently babble outside of interaction and that this seems to be important for phonological development may be more difficult to reconcile with the Linguistic hypothesis, however.[10] In this study, one baby babbled *more* frequently during solitary play than during dyadic interactions with the caregivers, two babbled 66% as frequently, and four babbled 50% as frequently during solitary play. A further two babies babbled almost equally frequently during solitary play and interaction. Vihman et al. (1985) also found that these babies'

---

[10] See also Oller et al.'s (2019) observation regarding babies' endogenously motivated protophone production during solitary play.

babble became more co-ordinated or 'speechlike' through repeated articulatory practice, regardless of whether it was produced solitarily or during interaction. If babble is a baby's attempt to produce language like an adult, we might expect to see more significant progress in 'speechiness' in response to situations where babies have access to rich information about language, modelled by a familiar adult in focused one-to-one interactions.

Naturalistic and theoretical work by Elbers and colleagues has also been interpreted as evidence that babble is underlyingly linguistic (Petitto et al., 2004). This interpretation, however, is not unquestionable. In two diary studies tracing the phonological development of one child, Thomas, from ages 6-15 ½ months and 15 ½ to 17 months, Elbers and Ton provide a wealth of evidence demonstrating significant continuity and interplay between babble and word production (Elbers, 1982; Elbers & Ton, 1985). Elbers (1982) describes how reduplicated (or *repetitive*) babble and a behaviour she terms *jargon babbling* – long babble sequences produced with somewhat language like prosody – seem to be necessary pre-requisites for Thomas' word acquisition. Elbers (1982) suggests that sounds made during reduplicated and jargon babble could acquire meaning and could thus become words. Elbers (1982) and Elbers and Ton (1985) also note that Thomas produced babble and first words overlappingly for a number of months, and that each could be seen to influence the form of the other. Thomas' production preferences in babble seemed to influence his selection of adult word forms to target, and he began to produce some of the sounds in his newly acquired words more frequently in his babble. For instance, when Thomas began to produced words containing /t, d/, the proportion of his babbles that contained /t, d/ rose from 15% to 40%.

Similarly, in a longitudinal cross-linguistic study involving 25 babies from 5 language backgrounds, de Boysson-Bardies and Vihman (1991) found that babies produced a higher proportion of stops, both in babble and words, in the months after they had begun to produce their first words compared with the time before this when they produced only babble. The authors point out that a purely motoric theory of babble might predict a higher proportion of stops in babble than in first words since stops represent a less difficult articulatory target. One explanation might be that, similarly to Thomas, these babies' new words exerted some influence over the consonants that they enjoyed producing. Another explanation may be that babies in this study tended to favour words containing less effortful consonants due to a trade-off between their still-developing neuromuscular control and their newly emerging cognitive resources for intentional communication, processing symbolic reference, and encoding arbitrary meaning. Evidence of mutual influence between babble and first words has been interpreted as support for the idea that babble and language are subserved by the same neural architecture.

Elbers (1982) also states in her concluding remarks that “[P]arental speech may have the important function of motivating the infant (‘triggering’ babbling) and of setting the goal (providing the infant with a general model of how speech should sound)”, but that Thomas’ phonological development showed no effect of language-related reinforcement from his caregivers. The speculative comment about ‘triggering’ and modelling has been cited by Petitto and colleagues (1991, 2002a, 2004) in support of the Linguistic hypothesis, the argument being that if babble is triggered by language exposure, then this must be mediated by some in-built capacity that makes babies sensitive to language and gives them some drive to do language. However, as previously discussed, it is possible that exposure to language along with other forms of adult behaviour may instead trigger a more domain-general drive in babies to behave like others around them (Donald, 1991). Additionally, if babies are innately sensitive to language input and babble represents their attempts to produce it, then the finding that babble does not show any influence from language-based caregiver reinforcement is difficult to explain.

In other parts of their paper, Elbers and Ton (1985) use more cautious language to talk about the commonalities between Thomas’ babble and his speech. The authors characterise babbling and taking as two separate but mutually influencing “speech ‘systems’”, underlining their perceived distinction between these behaviours (Elbers & Ton, 1985, p. 562). Whilst the authors use the word “goal” in their concluding remarks, they do specifically argue against teleological theories of language acquisition. Elbers (1982, p. 61) describes reduplicated babble as “a largely self-directed process of exploration during which the infant uses certain operating principles for constructing his own springboard into speech”. Elbers and Ton (1985), similarly, talk about how articulatory experience accumulates and is later adapted at the time when babies begin to ‘select’ words to target. So, whilst Petitto and colleagues interpret Elbers and Ton’s findings as evidence for the Linguistic hypothesis, it is not clear that these authors agree with this hypothesis. These findings do not preclude the possibility that babble may exist as a general, non-linguistic behaviour that, after some practice, becomes entrenched and flexible enough to be manipulated and extended to other contexts – namely, language – since language makes use of the same articulatory apparatus.

Ambient language influence and idiosyncratic production preferences

Evidence that there may be some connection between ambient language exposure and individual production preferences in babble has also been interpreted as support for the Linguistic hypothesis. Petitto and colleagues (1991, 2002a, 2004) suggest that babies' innate perceptual sensitivity to the patterns of language allows them to segment the speech stream and attempt to produce the resulting segments. It seems intuitive then that we should see some effect of ambient language on babies' individual production preferences.

Jakobson (1968) and Oller et al. (1976) found evidence that babies can sometimes be observed to produce consonants frequently in their babble that do not appear in their early words and that appear in only very few world languages. These individual production tendencies have been assumed to be endogenously motivated (Oller et al., 1976) and were previously cited in support of the Independence hypothesis (Jakobson, 1968), since babies were said to frequently produce sounds without a model in adult language. It is not clear how the Linguistic hypothesis would handle sounds like these. However, Vihman et al. (1985) also observed some such 'wild' babble consonants and noted that they sometimes appear in babies' early onomatopoeic words, in their productions of animal sounds modelled by caregivers, and in their imitations of environmental sounds that are not modelled by any caregiver (Laing & Bergelson, 2020) but not in their other early word forms. It is possible, then, that these sounds may in fact be modelled by adults in some specific (onomatopoeic) contexts which were not deemed to be 'linguistic' by previous researchers.

Work by de Boysson-Bardies and colleagues (1981, 1989, 1991, 1993) has demonstrated that ambient language may have some effect on which consonants babies settle on as Vocal Motor Schemes and how babies form vowels in their babble. In investigations of consonant production in babble, de Boysson-Bardies et al. (1981) and de Boysson-Bardies and Vihman (1991) found evidence that ambient language may influence consonant production. In a single case study of one French baby aged 18-20 months, de Boysson-Bardies et al. (1981) observed some common phonological tendencies between this baby and English acquiring babies: this baby produced few liquids, would reduce consonant clusters and devoice final consonants, and tended to produce more onset consonants than final consonants. However, this baby also produced a higher volume of nasalised vowels and fricatives than English infants in other studies, reflecting the higher volume of these sounds and phonotactic constraints in French than in other languages. de Boysson-Bardies and Vihman (1991) examined the place and manner of consonants produced by 20 babies acquiring English, French, Japanese, or Swedish (n=5 from each group) from the time when babble emerged until the time when babies were producing at least 25 distinct words within a 30-minute recording session. English and French babies were found to produce a higher proportion of labials than Japanese and Swedish babies both in babble and in their early words. English, French, and Swedish babies were also found to produce more stops than Japanese babies, and Japanese babies produced a higher proportion of fricatives in their babble and their first words than did babies in other language groups. The authors suggest that these production preferences reflect a combination of the respective frequencies of these sounds in the ambient language input, their visual salience (of labials, for example), and babies' ability to produce sounds and combinations of sounds. Labials are argued to represent particularly salient targets since lip articulations are visible externally. Whilst Swedish contains many initial fricative-stop clusters like /sk/, the authors suggest that the tendency for Swedish babies to produce more initial stops may be explained by babies' tendency to reduce clusters that they do not yet have the coordination to produce faithfully.

de Boysson-Bardies et al. (1989) analysed the front/backness (Formant 1) and the height (Formant 2) of 1047 vowels recorded from 20 ten-month-olds whose ambient languages were French, English, Cantonese, or Arabic. The study found significant variability across infants. Some produced a wide range of vowel sounds, whilst some produced a narrow range, but significantly larger differences in vowel production were seen between language groups than within them. The authors concluded that this finding could not be attributed to purely physiological or motoric factors, since all babies were equipped with the same articulatory apparatus, give or take minor individual differences in morphology. Furthermore, the formants of the vowels produced by babies in each group more closely resembled the vowels of their ambient language than those of other languages. de Boysson-Bardies et al. (1989) cite Werker and Tees’ (1984) finding that perceptual narrowing occurs around 9-10 months and comment that their own study provides evidence of *productive* narrowing. Perceptual narrowing is the experience-dependent process by which babies become more adept at perceiving frequently encountered sounds (ie., those in the ambient language), whilst becoming less sensitive to infrequently encountered sounds (i.e., those in other languages to which the baby is not exposed). Productive narrowing, likewise, is the experience-dependent process by which babies become increasingly proficient at producing favoured sounds (either those they enjoy producing (see McCune & Vihman, 1987) or perhaps those encountered in the ambient language) and less adept at producing disfavoured sounds (e.g., see Esling, 2012 on the prevalence of laryngeals in different language groups).

de Boysson-Bardies' (1993) study examined disyllabic babble utterances of 20 babies aged 10-12 months exposed to English, French, Yoruba, or Swedish (n=4 respectively). The phonotactic properties of these babies’ babble was compared to those of the disyllabic words of 18-month-olds acquiring the same language. Recall that the Motoric hypothesis predicts that front consonants should co-occur with low front vowels. de Boysson-Bardies (1993) found that different vowels co-occurred with labials compared with dentals, and that different vowels co-occurred with each of these consonant types in different language groups. For example, for babies acquiring English, Swedish, and French, dental consonants most often co-occurred with front vowels whilst for babies acquiring Yoruba, dental consonants most frequently co-occurred with central vowels. Further to this, these phonotactic preferences closely reflected those in the word form productions of the 18-month-olds acquiring the same language. de Boysson-Bardies (1993) interpreted these findings as evidence that ambient language exposure shapes babies’ babble production preferences. Arguing against the idea that the form of babble is determined by motoric constraints (as suggested by Frame/Content theory) Petitto et al. (2004, p. 46) cite de Boysson-Bardies (1993, p. 361): “…babies have a particular type of vocal apparatus at their disposal, but the constraints this apparatus puts on the production must be distinguished from the use to which babies put it”. However, Davis and MacNeilage (1995) have also responded to a second key finding from this study – that babies’ babble appeared to demonstrate more voluntary control over vowel height than over vowel front/backness, since vowel height was the dimension in which babies' vowels most closely resembled those of the target language. These authors argue that this finding supports their Frame/Content theory, which posits that the jaw is the primary articulator in babble, whilst the tongue maintains a near resting position. Davis and MacNeilage (1995) go on to critique de Boysson-Bardies (1993) and other studies seeking to establish whether ambient language influences babble, claiming that some of the observed trends may result from their small but diverse sample sizes of babies and observations.

McCune and Vihman (1987) found robust evidence that many babies will develop well-rehearsed articulatory routines or Vocal Motor Schemes, which allow them to produce a particular consonant or consonants voluntarily, reliably, and consistently in their babble and later, prominently, in their early words. This and later research by Vihman and colleagues (see Chapter 1) replicated this finding and found evidence to suggest that Vocal Motor Schemes may act as an *articulatory filter*, allowing babies to segment the speech stream and bootstrap into word acquisition (Keren-Portnoy et al., 2005; DePaolis et al., 2011; DePaolis et al., 2013; Majorano et al., 2014; Vihman et al., 2014).

The findings of these and other related studies strongly suggest that babble and language are, in fact, closely related. Whilst this is not necessarily problematic for more recent conceptualisations of the Motoric hypothesis, it is certainly strong evidence against the initially proposed Independence Hypothesis. Since Vocal Motor Schemes are heavily implicated in first word acquisition, if babble is linguistic – that is, if babble is babies attempting to produce language – then we should perhaps expect to find that language exposure determines which consonants babies produce the most. However, this more recent work has established that, whilst babies from different language backgrounds may seem to tend towards different Vocal Motor Schemes, the linguistic input that babies receive cannot be assumed to be the sole source of their Vocal Motor Schemes (McCune & Vihman, 2001; DePaolis et al., 2011; DePaolis et al., 2013; Majorano et al., 2014). Whilst some babies favour consonants or larger production patterns that are particularly frequent in their ambient language, other babies may favour consonants that a very infrequent (Elbers & Ton, 1985; DePaolis et al., 2011). Additionally, whilst the Infant Directed Speech of different caregivers has been found to show significantly similar phonological properties, babies' Vocal Motor Schemes have been found to show significant variability unrelated to these common properties (Vihman et al., 1994; DePaolis et al., 2011; DePaolis et al., 2013). Vocal Motor Schemes, then, seem to be significantly determined by exploration and enjoyment rather than purely by a drive to do language (Elbers & Ton, 1985; Elbers & Wijnen, 1992; Elbers, 2000)

Nature vs. nurture

The debate over the underlying nature of babble invokes the nature vs. nurture argument. Labelling babble as a mechanism *for* language acquisition and claiming that babble is triggered by a special brain-based sensitivity to language associates the Linguistic view more closely with the nature side of this argument (Petitto et al., 2004). A key weakness of this argument and of the Linguistic hypothesis is that it is necessarily teleological, whilst development in biological organisms is not. Studdert-Kennedy (1990, p. 17) writes that:

> "At each point in its development an organism is already complete, adapted and adapting, as best it can, to present conditions, internal and external. Just as earlier evolutionary forms existed for themselves, not for any later forms to which they might give rise, so the present form of a developing organism has its own present function. A child does not learn its first words so that it may later combine them into sentences. First words have their own economy."

By extension, it seems likely that babies do not begin to produce syllables in order to later combine them into larger units (words) or attach arbitrary meanings to them. Petitto and colleagues (1991, 2002a, 2004) do not discuss when or precisely how language input triggers babble but triggering of the type posited by proponents of the Linguistic hypothesis assumes that babies are capable of identifying a goal (adult language) and modifying their behaviour in order to move towards that goal. This speaks to a degree of intentionality and of articulatory planning and cognition of which we have little evidence to suggest that 6–8-month-old babies are capable. Whilst Petitto and colleagues (1991, 2002a, 2004) suggest that there may be some brain-based language capacity, they do not explain how this language capacity might allow babies to override the limitations presented by their as yet immature capacities for cognition, attention, and motor planning. The Linguistic hypothesis seemingly downplays the power and significance of accumulated articulatory experience in arguing that this cannot be sufficient to bring about babble emergence. However, as we have seen Chapter 1 and earlier in this chapter, accumulated productive and perceptual experience plays a critical role in phonological development.

Some human behaviours are more transparently under biological control, like bipedalism (Rosenberg & Trevathan, 2005; Trevathan, 2015). In humans, the position of the head and legs relative to the pelvis and spine makes walking on two legs more efficient than walking on four limbs, which is comparatively awkward and uncomfortable for the vast majority of adults. This is also specified in the brain: the motor cortical areas for the feet are more distal to the areas for the hands compared with the hind-and forelimb areas in quadrupedal mammals. Once they have developed sufficient neuromuscular and vestibular control, babies will begin to pull themselves to standing and begin to take their first steps (Thelen & Smith, 1994; Vihman, 2014). However, talking is not wholly like walking. Whereas the apparatus for walking evolved specifically for locomotion as their primary function, the articulatory apparatus involved in speech evolved for other purposes: the lungs for respiration, the lips, tongue, and jaw for feeding, and the larynx to protect the airway. The articulators of signed languages – the hands and facial muscles – evolved to serve a very wide range of functions including paalinguistic communication, feeding, self-stabilising or self-locomoting[11], manipulating the environment, and making tools (Thelen & Smith, 1994; Markze & Markze, 2000; Fragaszy & Crast, 2016). It cannot be assumed, therefore, that these organs nor that the parts of the brain that innervate them necessarily biologically predispose us to produce language.

Interestingly, there is some evidence that some of these organs have adapted *as a consequence* of language. Whilst the larynx sits high in the throat of non-human primates, the human larynx is descended (Vihman & DePaolis, 2000). This allows us to produce a much wider range of pitches, but also increases our risk of choking. Since evolutionary adaptations tend to occur only where they are fortuitous, this suggests that having access to

---

[11] For example, crawling in human infants, using a banister rail or walking aid in older children and adults, knuckle-walking in non-human primates

a wider range of pitches confers some social communicative advantage that outweighs the risk presented by suffocation. Humans are the only known animal with lips formed from thin, sensitive, hairless, mucosal skin (Bermejo-Fenoll et al., 2021). The *vermillion border* surrounding the lips may have become relatively dry and hairless as a result of language production over the last 200,000 years (Bermejo-Fenoll et al., 2021). Through prolonged exposure to air and frequent gentle friction between the lips when labial sounds were produced, Bermejo-Fenoll et al. (2021) suggest, this mucosal skin has been prompted to evolve, becoming resilient when dry and losing its hair follicles. In order for these changes to take place, humans or human ancestors have likely been producing linguistic or pre-linguistic vocalisations for some considerable time with these organs in their previous state. It seems, then, that babble and language 'makes do' with what resources the body has to offer and that these resources need not be pre-specified for language but may become slightly adapted over time. The same may be true of the brain

Evidence from laterality research

In efforts to address the nature vs. nurture debate, Holowka and Petitto (2002a, 2002b) have sought to establish whether the field of neuroscience can offer support for the Linguistic hypothesis. The authors examined asymmetries in babies' lip openings (i.e., orofacial asymmetry) when they produce babble compared with when the produce other vocal and facial gestures, aiming to establish whether there. Holowka and Petitto (2002a) conducted a cross-sectional study in which they analysed the lip openings – of 10 babies (5 French and 5 English) aged between 5-12 months whilst the babies were producing babble, other non-babble vocalisations, and smiles. Lip openings were rated +1, -1, or 0 to indicate whether a lip opening showed greater left-sided or right sided opening, or equal opening on each side. This study is of particular import to this thesis, as the present study aims to build on its findings. This and related research will be discussed in more depth in Chapter 3. Previous laterality research using behavioural, neuroimaging, and neuroanatomical methods has established that the left hemisphere of the brain is more active than the right when adults produce spoken language, and that the right hemisphere is more active that the left when adults smile (Graves et al., 1982; Borod et al., 1983; Gazzaniga & Smylie, 1983; Wyler et al., 1987; Borod et al., 1988; Graves & Landis, 1990).

Holowka and Petitto (2002a) applied Graves et al.'s (1982) method for measuring, rating, and comparing asymmetry in mouth opening (see 2.1) to a sample of still-frames extracted from video data of the 10 babies in their study. Babies were stated to have been visited at the exact point at which babble emerged. The results of the oral laterality analysis strongly suggest that the left hemisphere of the brain was more active than the right during babble. Since neuromotor control of the lower half of the face is contralateral, leftward asymmetry in babies' lip movements during smiles reflected greater involvement of the right brain hemisphere, whilst their non-babble vocalisations suggested much more variable hemispheric involvement. The authors interpret these findings as further support for the Linguistic hypothesis since babble and smile showed analogous differentiation to adult language and smiles. Holowka and Petitto (2002a) suggest that babble is the manifestation of the maturation of a brain-based capacity for language.

However, the wider picture may be less straightforward. A later cross-sectional study by Schuetze and Reid (2005) involving 41 babies at ages 12 (n=13), 18 (n=14) and 24 (n=14) months found evidence that the magnitude of babies' orofacial asymmetry in their emotional facial expressions increases over developmental time. This study failed, however, to replicate Holowka and Petitto's (2002a) finding of right hemisphere dominance during babies' smiles. Instead, babies in this study showed marked and increasing right hemisphere dominance for negative emotional facial expressions but closer to equal hemispheric involvement during smiles. One possible reason for this discrepancy may be the differences in sample size. Smiles typically emerge around 1½-3 months of age (Vihman, 2014) so by 12 months babies have already had many months of practice in producing them. Schuetze and Reid's (2005) findings suggest, though, that some brain specialisation in the form of lateralisation is still taking place well after the end of the first year of life. It seems likely then that a behaviour that is comparatively much newer and much less well-established – i.e., babble – should also show increasing lateralisation over time. Since Holowka and Petitto's (2002a) study is cross-sectional, it cannot speak to this question and further research is warranted to establish whether this is, indeed, the case. If babble is linguistic from the point of emergence, relying on a ready-made brain-based language capacity, then we might expect that babble should be stably left lateralised from emergence. However, since articulatory experience may play some role in the gradual lateralisation of emotional facial expression and perception (see Lindell et al., 2017) – a behaviour less controversially held to be universal and innate (see Ekman 1971, 1992, on the universality of smiling) – there seems to be good reason to question whether babble should show stable left hemisphere dominance from the time when it emerges.

Orofacial asymmetry research with non-human primates before and since the time of Holowka and Petitto's (2004) study has established that our closer and more distant cousins show a similar differentiation (Losin et al., 2008; Wallez et al., 2012). In chimpanzees, the right brain hemisphere has been found to show more activity during emotional calls, whilst the left brain hemisphere has been found to show greater activity during intentionally communicative attention-getting sounds (Losin et al., 2008; Wallez et al., 2012). This poses a particular problem for the suggestion that a brain-based language capacity that is special to humans is responsible for the left lateralisation that Holowka and Petitto (2002a) observed in babble. In light of these findings, it may be more fitting to say that Holowka and Petitto's findings represent support for the idea that *some of the capacities that support language* – rather than a specifically linguistic capacity – appear to be found predominantly in the left hemisphere. These may include capacities involved in vocalising for social reasons and cognitive capacities for understanding and manipulating social situations. This suggestion aligns more closely with the findings of orofacial asymmetry research with Common marmosets, who have been found to show left hemisphere dominance for social contact calls but right hemisphere dominance for emotional calls (Hook-Costigan & Rogers, 1998).

There are some key differences between the method for analysing orofacial asymmetry used by Holowka and Petitto's (2002a, 2002b) and the methods used by the other researchers whose work has been discussed in the section above (Hook-Costigan & Rogers, 1998; Fernandez-Carriba et al., 2022a, 2002b; Schuetze & Reid, 2005; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). One particularly important difference is that, in Holowka and Petitto's (2002a) study, still-frames showing babies' lip movements were rated into three categories by the researchers to indicate whether the left or right side of the mouth or neither opened more widely. The other studies outlined here have instead measured the area of the mouth opening by tracing the inner perimeter of the lips to calculate the area of the opening on each side of the mouth or *hemimouth*. Holowka and Petitto's (2002a) method, which was originally developed by Graves and colleagues' (1982, 1987, 1990) is less fine-grained and more vulnerable to subjective interpretation than the this alternative method of hemimouth area measurement (Fernandez-Carriba et al., 2002a, 2002b). Additionally, comparing the resulting areas of each hemimouth allows the researcher to examine the strength as well as the directionality by showing exactly how much wider or narrower the opening in one hemimouth is relative to the other. Further to this, analysing many still-frames for a single person or behaviour can elucidate whether people or behaviours show consistent degrees of left or right hemisphere dominance. This gives rise to much more nuanced and representative findings. The details and relative merits of these and other alternative methods for analysing orofacial asymmetry will be discussed in greater depth in Chapter 4.

Summary

The Linguistic hypothesis posits that human babies are born equipped with a special brain-based language capacity and that babble represents a baby's immature attempts at producing language, triggered when this brain-based capacity receives linguistic input. The regular timing and adult-like articulation of babble is taken as evidence that babies have a special ability to access the patterns of language. Unlike early conceptualisations of the Motoric hypothesis, the Linguistic hypothesis can account for the continuity that we observe between babble and later language. This hypothesis can also better explain how language exposure seems to exert some influence over the phonology of babble. However, the Linguistic hypothesis makes some claims that may not be psychologically or evolutionarily plausible.

Firstly, this hypothesis suggests that babble represents a seismic shift from previously existing vocal behaviours that can only be assumed to be motivated or 'triggered' by language exposure. However, as we have seen, articulatory practice and individual experience play a critical role in motivating babies to babble and equipping them with the phonological resources to do so, and the transition from early pre-babble vocalisations and into canonical babble is a fairly smooth and gradual one. The information that babies gain from language exposure does not straightforwardly determine how babies do babble or language. Under the Linguistic hypothesis, we might expect that babies should produce more babble in dyadic interactions and this babble should be more speechlike since babies are provided with a rich adult model of language. However, babies are approximately equally motivated to babble during interaction and during solitary exploratory play, and their babble becomes more ‘speechy’ (Oller, 1980, p.98) regardless of context, suggesting that self-monitoring and strengthening neuromuscular control are sufficient to increase its ‘speechiness’ (Oller, 1980, p.98). The idea of babble being ‘speechy’ may then be an artefact of adult perception, which simply equates to ‘coordinated’. Vihman et al. (1985, p. 440) write that “the famed division between babbling and language appears to reflect characteristics of adult perception and interpretation, rather than a developmental process manifested by the child”. It may be that viewing babble as a fundamentally linguistic behaviour is, similarly, a product of adult perception and interpretation. This view may be an artefact of adults (who already have language) viewing language as babies’ intended or predestined endpoint of their phonological journey and searching for the earliest point when the capacity for language can be said to be available.

A second key issue – and one that is not explicitly explored by proponents and opponents of the Linguistic hypothesis alike – is precisely what is meant by ‘linguistic’. Babble is different from speech and language in some critical ways. Whilst babble, like speech, can be described in both concrete phonological and metaphonological terms, it is produced without meaning or reference and without communicative intent (Oller, 1980; Vihman, 2014). Unlike babble, using language involves semantic, syntactic, and pragmatic processing, drawing on a range of different capacities and parts of the brain to the comparatively simple rhythmic phonological skills involved in babble (Bottini et al., 1994; Buchsbaum et al., 2001; Harpaz et al., 2009; Chang et al., 2010; Mesgarani et al., 2014; Sussman, 2015; Riès et al., 2016; Patel et al., 2018; Zhang et al., 2018;). Relatedly, babble is produced at a time when the developing brain has not yet reached full size or adult-like organisation – adult-like regionalisation is not reached until around age five and much specialisation and organisational change will continue to take place thereafter throughout childhood and adolescence (Mareschal et al., 2007). It may be that babble is neither wholly unprecedented in vocal-motoric development nor necessarily linguistic and that truly linguistic behaviour emerges only later in babies’ cognitive development when they begin to produce or combine words.

Thirdly, as we will see in Chapter 3 of this thesis, while there is a substantial quantity of research literature that examines how the brain supports language, none of this literature has identified any unique, specialised structure in the brain that could be considered to house a specific brain-based language capacity. In fact, fMRI (functional Magnetic Resonance Imaging) research by Häberling et al. (2016) and Häberling & Corballis (2016) has found evidence that language is supported by a network of different brain areas distributed across both hemispheres. What is more, the parts of the brain involved in the network supporting speech and language are involved in other functions like musical cognition, mathematical cognition, auditory perception, and general motor movement (Keenan et al., 2001; Griffiths & Warren, 2002; Nishitani et al., 2005; Meyer et al., 2012; Hodgson et al., 2021) as well as in language. In other words, these brain areas are *domain-general* rather than be specifically evolved to process language. MacNeilage (1998) has pointed out that for humans to possess a brain-based language capacity would require some sudden Chomskyan mutation (see also Crow, 2008). This would go against the widely accepted principal that evolutionary takes place incrementally over many generations i.e., the principle of descent with modification. Indeed, there evidence that the parts of the brain that support language processing actually long pre-date language and existed in our common ancestors with other, non-human primates and some other mammals (e.g., see Cantalupo & Hopkins, 2001; Losin et al., 2008; Wu and colleages, 2011a, 2011b; Wallez et al., 2012). This research will be discussed in greater detail in Chapter 3.

The Old Parts, New Machine hypothesis

This thesis puts forward an alternative hypothesis regarding the underlying nature of babble that is intended to address the gaps in the Motoric and Linguistic hypotheses. This hypothesis was outlined in Chapter 1 and argues that babble is a rhythmic-motor phonological behaviour that emerges endogenously out of a confluence of preceding vocal behaviours and skills, in the course of exploratory play. In this way, babble, it is argued here, is similar to other rhythmic motor behaviours seen during the first year of life. As with other vocal behaviours like cooing and turn-taking, babble production can sometimes be elicited by social interaction and is very likely involved in babies learning that the voice can be used for communication (Vihman, 2014). However, evidence suggests that babble, like other rhythmic behaviours, likely emerges spontaneously without the need for extrinsic triggering (see Thelen, 1979, 1981; Oller, 1980; Stark, 1980; Thelen and Smith, 1994, Ejiri & Masataka, 2001; Iverson et al., 2007, Oller et al., 2019) and, that, like other vocal behaviours that predate it, babble is equally likely to be produced during episodes of interaction or when a baby is contented and alone (Vihman et al., 1985; Oller et al., 2019).

The Old Parts, New Machine hypothesis posits that babble becomes directly relevant to, and influenced by, language only at the time when babies begin to recognise similarities between the sounds that they can make and the words that adults produce and are motivated to repeat or attend to particular sounds. However, it is argued that this recognition and repetition is underlain by a domain-general drive to behave like those around them – the same drive that motivates babies to clap or 'sing' when their caregivers do the same (Donald, 1991, Vihman & DePaolis, 2000). Before this can happen, babies must go through a kind of non-teleological 'practicing' period, during which they acquire and consolidate a range of articulatory and cognitive-attentional skills. Unlike the Linguistic hypothesis, this hypothesis argues that babble is not triggered when babies perceive language, nor is it produced by some brain-based capacity that evolved specifically to do language. Whilst it is widely accepted that there is significant continuity between the phonological form of babble and early words, it is argued here that this does not constitute evidence that the onset or the form of each baby's babble is determined by exposure to adult ambient language. In contrast to the Motoric hypothesis and Frame/Content theory, it is not suggested that babble emerges from ingestive cyclicities, nor that there is any discontinuity between babble and language. This hypothesis goes further in pointing out that there is significant *continuity* between babble and the behaviours that come *before* as well as after it.

This hypothesis – termed here the *Old Parts New Machine hypothesis* following the work of Bates et al. (1979, pp.; see also Bates, 1999, 2004) – draws on principles of evolution and development in dynamic biological systems in an attempt to provide a psychologically plausible account of what babble is, its role in phonological development, and how the human tendency to babble evolved. This hypothesis was discussed in depth in Chapter 1 with reference to phonological development in typically developing human babies. Further evidence for this hypothesis from cases of atypical development will be discussed in the following pages.

Evidence from atypical development

Atypical phonological development research can often give valuable insights into the specific roles of the different behaviours and capacities that are involved in babble and language and their emergence. For instance, in a study comparing oro-motor control in typically developing children and those with speech and language difficulties, Alcock (2006) found that oral motor skills predicted their overall language ability more strongly than their social and cognitive capacities. Twenty-one-to-twenty-four-month-olds who showed poor oro-motor control reliably showed poorer overall language skills, whilst those with stronger oro-motor control at this age showed a wider range of language abilities, more susceptible to influence from social and cognitive factors. Whilst this thesis focuses on laterality in typically developing babies, some research on atypical phonological development is discussed here to illustrate how this research has informed our understanding of the nature of babble. A comprehensive theory of development should be able to account for both typical and atypical developmental trajectories and this part of the chapter aims to demonstrate that the Old Parts New Machine hypothesis satisfies this condition. Recall that babble has been said to have three key pre-requisites: 1) a "biologically governed" capacity for [repetitive] rhythmic movement, and 2) sufficient neuromuscular control over phonation and movement of the articulators, which relies on 3) good quality auditory perceptual input and feedback (Koopmans-Van Beinum et al., 2001, p. 69). This section of the chapter will consider cases where access to one or more of these is disrupted or limited with a view to demonstrating the significance of these pre-requisites for the emergence of babble and later language acquisition.

Hearing impairment and babble

Koopmans-Van Beinum et al. (2001) proposed that good quality auditory input and feedback are critical pre-requisites for babble following a study investigating possible causes for delays in babble onset in hearing impaired babies. Oller and Eilers (1988), for example, traced the emergence of babble in 30 hearing and hearing-impaired babies and found evidence that the absence or low quality of auditory feedback may cause significant delays in babble and, in extreme cases, an absence of babble. All 21 hearing babies in this study began to babble between 6-10 months of age, whilst all nine hearing-impaired babies showed varying degrees of delay, with babble emerging between 11-25 months. Oller and Eilers (1988) remark on the fact that these nine babies did still receive some of the visual and auditory sensory information about speech and language available to hearing infants: they had some residual hearing, were provided with hearing aids, and received "speech-stimulation" interventions (p. 445). By contrast, a profoundly deaf child known to the authors who was born without cochleae and so had no exposure to speech had not yet begun to babble by age 3.

Koopmans-Van Beinum et al. (2001) investigated the phonetic and articulatory properties of the vocalisations of 12 babies aged 2-18 months. Hearing and hearing-impaired babies were found to arrive at key phonological milestones (like voluntary phonation, marginal babble, canonical babble) in the same order, but some delays were identified amongst hearing-impaired babies. All 6 hearing babies began to babble between 5 ½ and 7 ½ months. Five of the six hearing-impaired babies began to produce rhythmic voiced-unvoiced alternations during the period of study but did not begin to babble before data collection ended at 18 months. The one remaining hearing-impaired baby began to babble at 7 ½ months, and the authors suggest that this child may have had enough residual low frequency hearing to perceive auditory input from their caregivers and the wider world, and from their own output. Furthermore, when hearing-impaired babies did begin to coordinate articulatory gestures with voicing arrhythmically (i.e., when they began to produce marginal babble), these gestures tended to be velar, glottal, or pharyngeal gestures rather than labial, palatals, or alveolar gestures. These findings raise several important points. Firstly, that phonological milestones were attained in the same order in both groups of babies supports the idea that each phonological milestones in some way prepares babies to attain the next. Voiced, regular rhythmic open-closed alternations (canonical babble) seemingly cannot emerge before arrhythmic open-closed sequences with voicing (marginal babble), and these in turn cannot emerge before voluntarily controlled phonation. Secondly, the fact that hearing but not-hearing-impaired babies (with one exception) began to babble before 18 months suggests that the availability of auditory input and feedback plays a pivotal role in the integration of the capacities involved in babble. That some hearing-impaired babies in this study began producing sounds at the back but not the front of the vocal tract adds further weight to the idea that sensory feedback from self-produced sounds is important in phonological development. In the absence of good quality sensory information about language and what adults do with the voice, babies were still motivated to produce some consonantal sounds, and these tended to be the sounds that provided them with the most sensory feedback. These sounds may be more interesting for hearing-impaired babies as they create stronger vibrations in the throat when produced. By contrast, front and central consonants may give relatively little sensory feedback for hearing-impaired people aside from the tactile sensation generated by contact at the place of articulation.

Further research has reported similar findings. In research by Wallace and Yoshinaga-Itano (1998, cited by Menn, 1998, p. 524), their sample of 30 hearing-impaired 6-12-month-olds showed a preference for glottal consonants and, whilst eight babies did produce some oral stops, these were found in less than 10% of their babble. In more recent research, Persson et al. (2020) found that hearing-impaired 10-month-olds who received hearing amplification produced oral stops at a significantly lower rate than their hearing counterparts and produced a smaller range of consonants overall. Differences between hearing-impaired babies with hearing amplification and their hearing counterparts reduced by 18 months and the gap in development was found to close by 36 months. Exposure to auditory perception later during the first two to three years of life may then support hearing-impaired babies to 'catch-up' with their hearing peers.

As discussed in Chapter 1, the productive experience that babies gain during the early months or years shapes their transition into language. Earlier volubility and more sophistication or variety in babies consonantal repertoires are positively correlated with earlier developing language skills (Stoel-Gammon, 1992) demonstrating the predictive role of auditory and proprioceptive experience in phonological development. Observations that babble is delayed in hearing-impaired babies and sometimes absent in profoundly deaf babies further suggests that the tendency for babble to become entrenched is not simply determined by the maturation of some motoric or linguistic ability. Whilst babble emergence and incidental babble production may be purely motoric in the first instance/s of production, continued intentional production of babble is influenced by babies' ability to perceive the consequences of their own vocal behaviour. The fact that babies do produce and repeat some consonants regardless of whether they are able to perceive these in the adult input suggests that consonants emerge endogenously and the consonants that provide the most interesting or stimulating feedback are those that come to be repeated.

Down Syndrome and babble

Articulatory control is also vital for babble and speech emergence and babies with Down Syndrome face some particular challenges in this area (Nicole Whitworth and Karen Oxley, personal communication). Down Syndrome results when an extra copy of chromosome 21 develops by chance from the sperm or ovum during pregnancy, and this chromosomal change causes physiological and cognitive differences (NHS, 2019). Down Syndrome causes hypotonia (low muscle tone), and individuals sometimes have absent or superfluous facial musculature, which can obstruct or impede motor control over the jaw, lips, and tongue. Individuals with Down Syndrome also tend to have smaller airways and mandibles, and hard palates that are flatter (rather than arched), narrower, shorter, and positioned higher in the mouth compared with typically developing individuals (Stoel-Gammon, 2001). The tongue also tends to be larger relative to the size of the mouth and this, coupled with the hypotonia, can make articulation more effortful and less precise (Stoel-Gammon, 2001; Kent & Vorperian, 2013; Whitworth & Bray, 2015). The majority of individuals with Down Syndrome also experience some degree of hearing impairment (Stoel-Gammon, 2001, Whitworth & Bray, 2015). Immune system differences in Down Syndrome are frequently associated with increased vulnerability to chronic ear, nose, and throat infections and middle ear effusions like glue ear (Shott, 2006). Additionally, research has identified a number of different types of inner ear *dysplasia* (abnormal cell growth) in the inner ear bones, the ear canal, and the cochleae that affect a significant majority of individuals with Down Syndrome and can result in superfluous, absent, enlarged, or duplicated aural architecture or cavities and these things can compromise the auditory input and feedback that the brain receives (Blaser et al., 2006; Intrapiromkul et al., 2012).

Children with Down Syndrome tend to continue with babble for much longer than their typically developing peers – in some cases, up to 6-10 years of age (Nicole Whitworth[12] and Karen Oxley[13], personal communication). As in typically developing children, babble and words often co-occur but first words may not emerge until as late as 7 years of age (Stoel-Gammon, 2001; Nicole Whitworth and Karen Oxley, personal communication). Speech impairments and delays in babies, children, and adults with Down Syndrome have been found to stem not from cognitive deficits but from differences in physiology and neuromuscular control and are often aggravated by co-morbid dyspraxia and hearing impairment (Stoel-Gammon, 2001; Alcock, 2006; Whitworth & Bray, 2015). Some babies with Down Syndrome begin babbling within the same age ranges as their typically developing and hearing impaired peers and some much later (Stoel-Gammon, 2001, Whitworth & Bray, 2015). The majority of individuals with Down Syndrome require much more vocal practice than their typically developing peers in order to develop the phonological resources for speech, in particular, the requisite neuromuscular control and the ability to link acoustic signals in the environment to in repertoire articulatory motor routines (Stoel-Gammon, 2001; Whitworth, personal communication). Whitworth and Bray (2015, p. 4) found evidence that these physiological differences cause Down Syndrome babies' babble to be more unsystematic. The authors observe that babies and children with Down Syndrome tend to produce "fuzzy, unstable representations of somatosensory and auditory targets, which remain in a state of constant flux" throughout their phonological development. They argue that this is the cause of their delayed language acquisition. So, Down Syndrome research further demonstrates the importance of articulatory practice and developing neuromuscular control for phonological development.

The individuals with Down Syndrome involved in the research described here are typically provided with a good deal of intensive, targeted speech and language-based interventions and therapies in order to support their development (Stoel-Gammon, 2001, Kent & Vorperian, 2013; Whitworth & Bray, 2015). In spite of this additional language exposure and

---

[12] Head of Subject for Languages at Leeds Beckett University, formerly Lecturer in Speech and Language Therapy

[13] Registered Speech and Language Therapist with expertise in paediatric speech and language therapy and Autism Spectrum Conditions. Previously employed in the NHS and privately (Away With Words, West Yorkshire).

the cognitive drive to behave like those around them, babble and words only emerge when the child has the requisite motor control (Stoel-Gammon, 2001, Kent & Vorperian, 2013; Whitworth & Bray, 2015). Interestingly, babies and children with Down Syndrome are able to feed for many months and sometimes years before babble and first words emerging, suggesting that the rhythmic jaw oscillations involved in eating do not spread into vocal behaviour as suggested under Frame/Content theory (Davis & MacNeilage, 1995; MacNeilage, 1998, 2008).

William Syndrome and babble

Rhythmic ability is also instrumental in the emergence of babble and speech. Individuals with William Syndrome exhibit a number of physiological and cognitive differences from typically developing individuals, which give rise to differences in motor, social, linguistic, and rhythmic development (Lenhoff et al., 1997; Masataka, 2001; Mervis & Becerra, 2007; William Syndrome Foundation, n.d.). Whilst William Syndrome children and adults characteristically have very large vocabularies populated by complex and low frequency words, they often show lower processing ability with other aspects of language like grammar and pragmatics (Mervis & Becerra, 2007). Additionally, the first words usually emerge later than in typically developing babies, at around 18-24 months (Lenhoff et al., 1997; Masataka, 2001). This may result from cognitive differences, since William Syndrome is usually associated with learning difficulties and fairly low IQ (William Syndrome Foundation, n.d.), but it seems likely that differences in rhythm and gross and fine motor coordination may also play some role here.

Babies with William Syndrome also typically begin to walk later, towards the end of the second year; speech sound articulation remains effortful and imprecise even after the emergence of first words; and pencil control takes several more years to master than in typically developing children (Lenhoff et al., 1997; Masataka, 2001). When babies with William Syndrome do learn to walk, they often balance on the balls of their feet and their movements have been described as “awkward” (Masataka, 2001, p. 160). Babble research with William Syndrome babies also supports this suggestion. Masataka (2001) studied eight babies with Williams Syndrome from 6-30 months of age and observed a delayed onset for babble, first words, and other rhythmic and motor behaviours including rolling, sitting, standing, walking, and hand-banging. Masataka (2001) identified significant correlations between the onset of hand-banging and a) babble and b) first words, as well as the expected correlation between babble and first words, mirroring findings in typically developing babies (Thelen, 1979, 1981; Ejiri, 1998; Iverson & Thelen, 1999; Ejiri & Masataka, 2001; Iverson & Fagan, 2004; Iverson et al., 2007; Fagan, 2009). Some variability was noted across babies with respect to age of onset: babies began to babble between 15-21 months and began to produce their first words between 18-24 months. Masataka (2001) concludes from these findings that rhythmic development is a control parameter for babble and, by extension, for language development.

Like babies and children with Down Syndrome, those with William Syndrome typically receive targeted speech and language therapy and yet, as in Down Syndrome, babble and language do not emerge in response to this rich language exposure but instead in direct response to an endogenously originating control parameter: in this case, rhythmic ability (Masataka, 2001). Once again, the stable rhythmic jaw oscillations developed in feeding behaviour do not seem to equip William Syndrome children adequately to master canonical syllable production. The ability to manage the rhythms of language seem to emerge from elsewhere – and not "fully formed" (MacNeilage & Davis, 2001, p. 179)- and continue to become more precise and adult-like through practice during babble and first speech. Interestingly, this latter insight from William Syndrome research suggests that fully adult-like consonant-vowel articulation, whilst it may be characteristic of typical babble and speech, is not actually a strictly necessary pre-condition for speech – articulation simply needs to be adult-like and precise *enough* to be understood by others.

Tracheostomy and babble

Alongside cases of typical development, the role of phonatory and articulatory experience has also been illustrated through cases studies of early tracheostomy. Tracheostomy involves the incision and cannulation of the trachea below the larynx and is undertaken in cases where individuals are unable to breath without assistance. In babies, this may be because the airway is underdeveloped, obstructed, injured, or because fluid has settled there (NHS, 2017). Whilst a small amount of air still flows in and out of the body through the larynx, whilst the cannula is in place, this volume is not sufficient for modal voicing, meaning that tracheostomised babies cannot produce the same quality of vocal exploratory play as typically developing babies (Locke and Pearson, 1990). These babies are termed ‘aphonic’ in the literature and tracheostomy has been found to disrupt babble emergence and the transition into language after the trachea and neck are surgically repaired (see Locke & Pearson, 1990; Bleile et al., 1993; Bohm et al., 2010).

Locke and Pearson (1990) conducted a single case study of one baby, tracheostomised for 15 months from 5-20 months of age. After decannulation, the baby was observed to produce a very low rate of babble – around 90% less than typically developing babies. The babble that this baby did produce showed an apparent preference for labial sounds. Bleile et al. (1993) discuss the more limited progress in speech and language development made by another baby following decannulation at 28 months. This baby had been tracheostomised for 27 months beginning at 1 month of age. Whilst this baby had ‘caught up’ developmentally by age 4, during the initial months after decannulation, they showed a developmental gap of around a year between their receptive and productive capacities. This child also showed a low rate of babble containing a narrow range of consonants, with a preference for labial sounds (consonants involving closure of the lips like /b/ or /m/), like the baby in Locke and Pearson’s (1990) study. Additionally, this baby did not begin to use voicing voluntarily nor to coordinate this with intentional lip and jaw movements for at least the first five days following decannulation. Locke and Pearson (1990) and Bleile et al., (1993) have suggested that the observed preference for producing labials may result from the relative ease of producing labial, requiring simple opening and closure of the jaw, or from the availability of visual information about labials in caregiver speech.

Bohm et al.'s (2010) multiple case study of seven babies tracheostomised at 2 months and decannulated between 13-34 months of age made similar findings. Bohm et al. (2010) outline the emergence of different vocal behaviours relative to the time of decannulation (see Table 2). Babies in this study were observed to go through a brief period of producing vegetative sounds characteristic of those produced by typically developing babies in the very early months some time before babble emerges. Some weeks following this, babble emerged and, later, first words. It is not stated but it is reasonable to predict that these children went through some or all of the phases outlined in Chapter 1 following this post-cannulation vegetative period and before babble could emerge. As noted in Chapter 1, the term ‘vegetative’ is sometimes extended to include all pre-babble vocalisations.

Babies in all of these studies were reported to have begun to produce small numbers of words several weeks or months after decannulation, but these were reportedly limited in number and sometimes in their phonological form (Locke & Pearson, 1990; Bleile et al., 1993; Bohm et al., 2010). However, Bleile et al. (1993) comment that the productive repertoires of previously tracheotomised babies tend to reconcile largely with their receptive language repertoires by age 5. The productive-experiential deficit that results from aphonia and the subsequent delay in a baby’s vocal development and later word acquisition, again, demonstrates the value of auditory and proprioceptive experience for phonological development. The fact that babble was observed to occur in all previously tracheostomised babies before they acquired words adds weight to the argument that babble is critical for language acquisition. Most babies with tracheostomies continue to feed normally – only the trachea is incised and not the oesophagus (NHS, 2017). In spite of this, decannulated babies go through a period of articulatory practice before babble emerges, suggesting that jaw oscillations from feeding do not give rise to babble once voicing becomes possible after reparative surgery. Rather, the requisite jaw motions more likely arise through articulatory practice gained during this so called ‘vegetative’ phase.

Time post-decannulation	Behaviour emerging
2-4 weeks	vegetative sounds
2-12 weeks	canonical babble
3-14 weeks	first words

Table 2 Phonological behaviours and their respective times of emergence following decannulation (Bohm et al., 2010).

Autism Spectrum Conditions and babble

Children with Autism Spectrum Conditions have being found to show a higher rate of oral motor skill difficulties than neurotypical children as well as some language differences

(Alcock, 2006). Autism Spectrums Conditions are a range of developmental and social conditions, caused by neurological differences, which are associated with differences or impairments in interests and habits, social communication, emotional processing and regulation, executive function, access to education, employment, and civic life (NIMH, 2022).

Nine-to-eighteen-month-olds who were later diagnosed with Autism Spectrum Conditions have been found to be less voluble overall, produce a lower proportion of babble, and show delayed onset of babble when compared with their typically developing peers (Patten et al., 2014; Warlaumont et al., 2014). A possible reason for this is that Autism Spectrum Conditions may impair the caregiver-child "social feedback loop" (Warlaumont et al., 2014, p. 1314): the way that cyclical contingent vocal and socio-emotional interactions between baby and caregiver contribute to phonological and communicative development, and later language acquisition (Meins, 1998; Gros-Louis et al., 2006; Gros-Louis et al., 2014; Patten et al., 2014; Warlaumont et al., 2014; c.f. the production-perception feedback loop).

Contingent social feedback reinforces babble and other infant vocal behaviours. When babies babble, they are 'rewarded' with contingent attention and interaction from their caregivers, and the phonological form of this interaction often closely reflects that of the baby's babble (Gros-Louis et al., 2006; Gros-Louis et al., 2014; Patten et al., 2014; Warlaumont et al., 2014). Contingent feedback may involve speech or may involve adults repeating the baby's babble or even facial expressions and other movements back to them (Gros-Louis et al., 2006; Gros-Louis et al., 2014; Warlaumont et al., 2014). Some caregivers repeat their babies' vocalisations back to them or engage in 'conversations' with their babies using nonsense words, 'baby' talk, and/or infant directed speech (e.g., "oh really? babababa? and then what happened?"). Caregivers'' contingent vocal feedback tends to be produced with mature resonance and transition timing and it is suggested to a) provide babies with motivation to repeat the vocalisations that got them the attention, and b) provide babies with an articulatory model to approximate (Warlaumont et al., 2014). Contingent feedback has been found to be associated with more frequent production of babble and with more sophisticated articulatory control.

Among babies later diagnosed with Autism Spectrum Conditions, the social feedback loop may be disrupted or limited in three ways (Warlaumont et al., 2014). Firstly, these babies' exposure to contingent feedback may be delayed due to the delayed onset of babble. Secondly, feedback may also be limited in quantity since lower volubility generates fewer opportunities for contingent feedback. Thirdly, since Autism Spectrum Conditions are held to primarily affect social processing and functioning, babies with Autism Spectrum Conditions may be less responsive to contingent feedback or may find it less rewarding or motivational than their typically developing peers. If the input to the social feedback loop (the baby's behaviour) is disrupted or limited, then its output (the adult response) will also necessarily be constrained. This may mean fewer contingent interactions that might have the power to influence the baby's future behaviour and so their development. If babies' interest is not captured by the contingent social feedback, then a social feedback loop is unlikely to present a strong motivator for development and change.

Such cases illustrate the instrumental role played by a baby's domain-general interactions with their environment and their propensity to attend – or not – to the consequences of these interactions. The Motoric hypothesis, focusing, as it does, on the mechanics of babble, offers no account for how social processing might be involved in babble. The Linguistic hypothesis does take feedback and interaction into account but focuses specifically on linguistic information rather than social information more generally. Whilst the medium of the social feedback described here may often be language, this is not exclusively the case since caregivers' singing and imitations of their baby's vocalisations are also frequently observed during caregiver-child dyadic interactions (Gros-Louis et al., 2006; Gros-Louis et al., 2014; Warlaumont et al., 2014). It also does not necessarily follow that a baby's motivation and ability to approximate adult models in these scenarios is inherently linguistic since babies can be seen to repeat or approximate a range of different language-based and non-language-based sounds that their caregivers make – including their own babble. Again, the drive to repeat the behaviour seems likely more to do with social factors like securing some rewarding attention or the desire to behave like the caregivers

Summary

In Chapter 1 we saw how the hypothesis put forward in this thesis – the Old Parts New Machine Hypothesis – can account for babble emergence, for babies' attainment of phonological milestones within babble, and for how babies begin to make the transition into first word acquisition. In this section, we have seen how this hypothesis can also offer a better account for atypicalities in babble development than the Motoric or Linguistic hypotheses. These and other cases of atypical phonological development can offer valuable insights into the roles of the different capacities that come to be bound together in babble and language acquisition. Delayed onset of babble and low rates of production in hearing-impaired babies and those later diagnosed with Autism Spectrum Conditions highlight the role of the social and perceptual-productive feedback loops. Atypical babble in tracheostomised babies and those with Down Syndrome and William Syndrome illustrates the role of vocal articulatory practice and rhythmic ability. The findings of the research discussed here add further support to the hypothesis that babble emerges endogenously as a confluence of behaviours and capacities and later, after the accumulation of some considerable articulatory experience, comes to be relevant to another behaviour that is cognitively distinct from babble but uses overlapping articulatory apparatus - language. The universality of the impulse to babble suggests that it may be partly intrinsically motivated by some maturation in the central nervous system. However, there seems to be little evidence to suggest that maturation or triggering of an innate brain-based language capacity is responsible for babble. The influential role of caregiver interaction and auditory and proprioceptive input and feedback suggests that babble is also partly an experience-dependent social dynamic process and not a static or even stable form of behaviour arising through maturation.

The evidence outlined here and in Chapter 1 speaks to the motivation for working to understand babble. Furthering our understanding of the diverse capacities underlying babble can help to establish how typical and atypical babble diverge. Identifying points of divergence may facilitate earlier diagnosis of children who are likely to face language related difficulties later in life and can inform how we construct interventions and therapies for individuals affected. This is of value since it is widely observed that earlier intervention is typically associated with better outcomes. Recall that differences in oro-motor control in typically and atypically developing toddlers predict language and communication outcomes later in life (Alcock, 2006). Phonological skill in babble and the first words, then, seems to be of particular importance for language development. This may relate to the idea that having stable and entrenched ways of producing the sounds of speech ‘frees up’ cognitive and attentional resources that can then be used to combine sounds or process symbolic reference. By contrast, if oro-motor control is poor by the time the social and cognitive pre-requisites for language come into play, the transition into language may be slower or more effortful, or even disrupted.

The evidence discussed so far in this thesis also suggests that the rhythmic and motor capacities in the developing brain play a critical role in babble emergence and production. This being the case, the development of these parts of the brain and the behaviours for which they are responsible merit further attention from research. The present study aims to address this gap in the existing research by examining babies’ laterality during babble and other behaviours and asking whether any changes in lateralisation can be seen with time and articulatory experience. Once we have developed a more detailed picture of the laterality of babble in typically developing babies, future research may seek to establish whether any differences exist between typically and atypically developing babies with regard to laterality. The present study aims to lay the groundwork for future studies involving neurodivergent and atypically developing babies and children.

If differences in laterality are found to be associated with neurodivergence, for example, this knowledge may help us to identify and support babies who are at risk of conditions like dyslexia, dyspraxia, or Autism Spectrum Conditions. Neurodivergence is typically diagnosed late in childhood when children reach school and sometimes not until much later in adolescence or adulthood, by which time language difficulties may already have begun to affect quality of life. Earlier identification and interventions could reduce differences in language acquisition between the infant and their peers and reduce the impact of some of the social and communicative difficulties, and the self-esteem and academic attainment issues faced by children with these conditions by giving them the tools to better express themselves and interact with others.

In conditions which can be diagnosed earlier, understanding differences in lateralisation for babble could inform how we design speech and language interventions. Children with Down Syndrome and William Syndrome often receive physical therapy and speech and language therapy to support their motor development and their neuromuscular control. Laterality research may contribute to an understanding of some other ways in which these children's development can be supported.

Concluding remarks

A satisfactory and comprehensive theory of babble should explain what babble is and account for how it emerges and how it shapes phonological development. The Motoric hypothesis offered a conceptualisation of babble and its emergence that is broadly consistent with principles of evolution. The Linguistic hypothesis offers an account of how babies make the transition from babble into language. However, it could be argued that the Linguistic hypothesis, like other Nativist theories, places too much emphasis on the adult state of language as the endpoint of phonological development and attempts to reverse engineer how babies arrive there. Whilst reverse engineering may seem to be our only recourse in theorising about language phylogeny, other approaches are available to us when considering the ontogeny of language.

Rather than assuming the adult state as a basis and taking a top-down approach to thinking about phonological development, we can instead observe bottom-up what children actually do during their journey into language and how they come to arrive at the adult state and draw conclusions from this. Understanding development as a series of cyclical adaptive changes occurring in response to interaction between the baby, their own body, and their environment resulting in ever more complex behaviours and skills allows us to consider the multiple and various factors that influences phonological development and how these are intricately interconnected.

This approach also eliminates the need to posit that humans possess some special brain-based language capacity for which we have no neurological evidence, or that babies under the age of one are capable of identifying and working towards long-term goals and following phonological and phonotactic 'rules'. Instead, we have behavioural tendencies, which are shaped by a baby's experiences, and which come to constrain the experiences that they will go on to have in the future. The next chapter will explore evidence from studies of the brain and reflect on how their findings speak to the hypothesis proposed here.

Chapter 3 Lateralisation and the brain

This chapter opens by introducing laterality and lateralisation. Research concerning the laterality of language-related functions in humans is introduced to provide context for and evaluate the hypothesis that babble may be fundamentally linguistic. Following this, research concerning the lateralisation of language-related and non-linguistic functions during development and evolution is reviewed. The research reviewed here calls into question whether the capacities involved in babble can be considered to be truly linguistic, but also raises further problems for purely motoric accounts of babble emergence and development. Instead, the findings of this research promote a view of babble as a dynamic confluence of domain-general behaviours and capacities – a view more closely aligned with the Old Parts, New Machine hypothesis of babble. Methods for analysing laterality through orofacial asymmetry are also reviewed.

"The fact that humans have formal language makes some people believe that we humans are unique, or special. And whilst we may be special, processing complex auditory signals on the left side of the brain may not be what makes us that way."

(Poremba, 2006, p. 87)

Introduction

In neuroscience, laterality refers to the dominance of one hemisphere of the cerebrum or cerebellum over the other in performing a given function or set of functions (Corballis, 2008; Vallortigara et al., 1999). Lateralisation refers to the processes by which this hemispheric dominance comes about – that is, the process by which structures and functions become localised to the hemispheres and regions of the brain that typically support them

(Vallortigara, 1999). Some functions, like motor control, become lateralised through biologically-determined structural developmental processes during gestation (see Mareschal et al., 2007). Other functions, like processing of pitch, are recruited to one hemisphere or the other through experience-dependent developmental processes (see Minagawa-Kawai et al., 2011). Most behaviours are supported by activity in networks of brain areas distributed across both hemispheres and so laterality does not mean exclusive control by one hemisphere or the other. It was initially believed that cerebral asymmetry was unique to humans and language, but a growing body of laterality research has found evidence of hemispheric specialisation for various functions throughout the *Bilateria* (Vallortigara et al., 1999; Poremba, 2006; Corballis, 2008; Lindell, 2013).

Lateralisation is theorised to be an adaptive response to evolutionary pressures associated with survival (Vallortigara et al., 1999; Corballis, 2008). Vallortigara et al. (1999) write that there is a "functional incompatibility between the logical demands associated with very basic cognitive functions" (p. 164). When an organism encounters some object, phenomenon, or other organism in their environment, they must implicitly i) compare the present experience with past experiences; ii) use salient and appropriate cues to categorise parts of the present experience (e.g., as food vs. non-food, safe vs. threatening); and iii) represent the present experience and its salient constituent parts in memory for future comparison (Vallortigara et al., 1999). These two types of processing – the global and holistic vs. the sequential and segmental – require different types of cognitive computation (Vallortigara et al., 1999; Streri & de Hevia, 2014; Sussman, 2015). Recruiting each type of computation to a different hemisphere allows the organism to perform both types of processing simultaneously, decreasing the time taken to respond appropriately (Lindell, 2013).

However, whilst optimising efficiency of processing, lateralisation may involve compromises (Corballis, 2009). The natural world is not structured around any implicit lateral bias, so animals with a bilaterally symmetrical body plan[14] can adapt their interactions with their

---

[14] (that is, animals who have a right and left side, which are a mirror image of one another e.g., two arms, two eyes etc.)

environment on-line with greater flexibility than those with asymmetric body plans[15] (Corballis, 2008, 2009, 2015, 2017). Duplicated ears and eyes give us directional hearing and depth perception. Duplicated organs in the immune, respiratory, renal, and reproductive systems can preserve health, life, and procreation following unilateral organ damage; duplicated hands allow us to perform two discrete or complementary actions simultaneously; near duplication of the hemispheres of the brain allows us to innervate both sides of the body with near equal acuity (Vallortigara, 1999; Corballis, 2008). However, in some cases, the evolutionary advantages of functional asymmetries can outweigh the associated risks. Especially complex biological and cognitive functions that consume a great deal of space and processing power in the body are typically represented only once, giving rise to internal asymmetries (Lindell, 2013). Duplicating complex cognitive capacities might preserve functionality after injury or allow us think and socialise in more creative or flexible ways, but would require a larger cranium, more processing time, and more energy; would be more susceptible to cross-hemisphere interference; and may not confer any advantage outweighing these costs (Lindell, 2013; Corballis, 2015, 2017). Primate cranial volume is constrained by our reproductive anatomy: foetal head size is already “not uncommon[ly]” associated with birth complications and fatalities across species, from marmosets to humans (Rosenberg & Trevathan, 2005, p. 161; Trevathan, 2015). Primates have adapted to minimise these risks by being born prematurely relative to other animals, by evolving an intricately folded cortical morphology, and by giving birth in proximity to others for social and survival-related reasons (Michel, 2003; Rosenberg & Trevathan, 2005; Trevathan, 2015; Corballis, 2017; Demuru et al., 2018). Lateralisation of complex cognitive capacities may represent a fourth adaptation, allowing primates to live complex mental and social lives in spite of constraints on cranial and cerebral size.

Species-level patterns of brain asymmetry may also contribute to maintaining social cohesion and co-operation by promoting similar patterns of behaviour amongst individuals, though this may make some animals disadvantageously predictable to their predators or prey e.g. shoal behaviour in fish and pack behaviour in hyenas (Vallortigara et al., 1999). Differences in cerebral asymmetry at the level of the individual may serve to reduce this

---

[15] C.f. Organisms whose body plan does have a lateral bias like flatfish, scale-eating cichlids, and fiddler crabs are inherently more constrained in their sensory perception, feeding behaviour, and manoeuvrability than bilaterally organised organisms.

predictability and distinguish individuals from one another (Vallortigara et al., 1999; Corballis, 2009). Solitary species, who neither benefit from nor are constrained by the safety in numbers given by herd membership and so for whom unpredictability represents a significant advantage, show greater individual variability in brain organisation than social species (Vallortigara et al., 1999). Individual-level variability in laterality may also serve to reduce competition amongst members of solitary and social species alike, predisposing individuals towards different preferences and patterns in survival-related behaviours like feeding or reproducing (see Hori, 1993, for an extreme example of lateralisation in fish).

Its ubiquity suggests that laterality is a particularly ancient trait, governed, at least partly, by genetic mechanisms (Vallortigara, 1999; Corballis, 2008). One idea is that lateralisation may come about through active suppression of functionality in one hemisphere by areas in the other via the corpus callosum (Annett, 1996; Lindell, 2006; Poremba, 2006; Sussman, 2015). Other complementary ideas are that different neurons are predisposed by their biochemical make-up and the distinct rhythms at which they oscillate to process the structure of sensory information in different ways (Johnson, 2000; Mareschal et al., 2007; Morillon et al., 2010). During foetal and early-life development, neurons migrate through layers of the cortex towards higher concentrations of neurons sharing similar biases and rhythms and away from neurons with contrasting properties in order to maximise the efficiency of information transmission (Johnson, 2000; Mareschal et al., 2007; Sussman, 2015). The neurons that migrate to the left hemisphere are typically more proficient at segmenting sequentially organised information (like the speech stream), and retrieving hierarchically structured information from memory (Lindell, 2006; Poremba, 2006; Sussman, 2015; Riès et al., 2016). Neurons in the right hemisphere are typically more adept at processing emotional and visuospatial information holistically as gestalts and recognising patterns within or across gestalts (Lindell, 2006; Poremba, 2006; Sussman, 2015; Riès et al., 2016). That is not necessarily to say that each hemisphere is incapable of performing types of processing typically lateralised to the other, though their performance of these functions may in some cases be slower, more effortful, or more limited (Sussman, 2015). Cross-hemispheric suppression and neurochemical and oscillatory differences may contribute to explanations of how the brain achieves global and sequential processing simultaneously and in partially overlapping networks of neurons, whilst avoiding cross-task interference.

Whilst genetics undoubtedly plays an important role in shaping the brain, it may be more responsible for the fact of laterality than for the direction in which it is expressed in certain functions (Vallortigara et al., 1999; Corballis, 2009). Experience-dependent processes of cerebral specialisation are also involved in determining the developmental trajectory and direction of asymmetry in functions supporting complex behaviours like language, certainly from the time of birth but perhaps also pre-natally (Previc, 1991; Corballis, 2009; Tierney & Nelson, 2009; Bishop, 2013; Lindell, 2016). Lateralisation and regionalisation may then be governed by two interacting guiding principles of *functional specialisation* and *functional integration* (Serrien et al., 2006). The former refers to the particular computational aptitude of each neuron, and the second to the multiple and various ways in which neurons are connected, are coactivated, and interact – that is, the way they are integrated – in distinct networks to support specific functions. In order for the utility and aptitudes of each neuron and its connectivity to be established, the organism must make use of them (or not). Fortuitous neurons and synapses are preserved, sometimes remodelled, and strengthened, whilst neurons and synapses that do not contribute efficiently to any function decay and are pruned (Mareschal et al., 2007; Petanjek et al., 2011). Functional integration is, in every instance, influenced and constrained by the successive integrations that have gone before (Thelen & Smith, 1994; Vallortigara et al., 1999). Information with similar properties is usually stored or processed more proximally in the brain than information with very different properties (Buchsbaum et al., 2001; Chang et al., 2010; Mesgarani et al., 2014). Creating representations of wholes or parts of experiences and the organism's responses to them thereby results in recruitment and localisation of functions to specific areas within one hemisphere or sometimes both hemispheres.

This chapter concerns the lateralisation and localisation of vocal and communicative functions.

Laterality of vocal and communicative functions

For the majority of people, structures in the left cerebral hemisphere are responsible for more of the neural activity that supports language processing than structures in the right (e.g., see Graves et al., 1982; Häberling & Corballis, 2016; Häberling et al., 2016). This proportion is higher amongst right-handed people than left-handed people, at up to 95% and up to 70% respectively (Rasmussen & Milner, 1977; Graves et al., 1982; Knecht et al., 2000; Nenert et al., 2017; Van Der Haegen et al., 2018), but this species-level tendency towards left laterality is robust. Amongst the remaining 5-30% of people, some exhibit right lateralised of language-related functions, and, in a smaller proportion, language is represented bilaterally (Rasmussen & Milner, 1977). Early work by Dax, Broca, Wernicke, and Hippocrates (cited in Lindell, 2006 and Corballis, 2015) observed that lesions to certain areas of the left hemisphere resulting from illness or injury are associated with dysphasia or aphasia and sometimes hemiplegia or paraplegia of the right side of the body. Since these initial observations were made, several overlapping research fields have emerged examining the laterality of language; its ontogeny and phylogeny; and its relatedness to emotional and motoric functioning. This chapters explores the relevance of research from these fields to the present thesis and the hypothesis of phonological development that it puts forward.

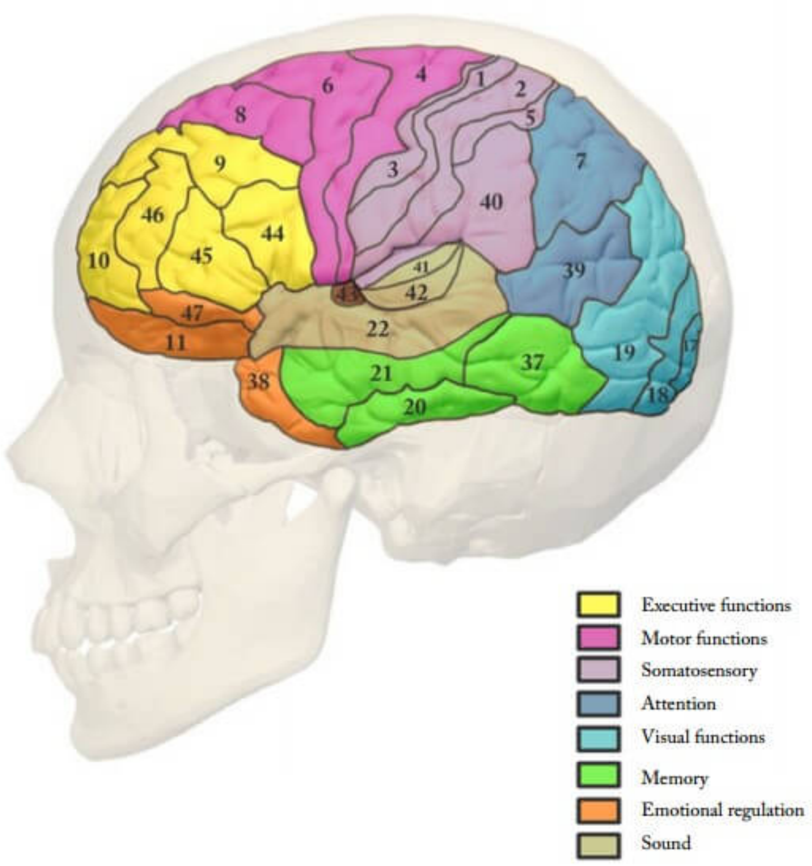


Figure 1 Map of Brodmann areas in the human brain. This image is reproduced from https://epomedicine.com/medical-students/brodmann-areas-lesions/

The classical left (or dominant) language areas include Broca's area (the pars opercularis and pars triangularis on the inferior frontal gyrus, Brodmann areas (BA) 44 and 45 respectively) and Wernicke's area (the posterior superior temporal gyrus, BA 22, which houses the planum temporale) (see Figure 1) (Müller & Basho, 2004; Harpaz et al., 2009). Other left-hemispheric areas involved in language include the medial and anterior superior temporal gyrus (BA 41, 42), superior temporal sulcus, supplementary motor cortex, medial prefrontal gyrus, precentral gyrus, anterior fusiform gyrus, anterior cingulate gyrus, angular gyrus, orbitofrontal cortex, frontal eye field, lateral temporal pole, temporo-occipito-parietal junction, and parts of the cerebellum (Leiner et al., 1993; Dronkers, 1996; Dum & Strick, 2003; Spitsyna et al., 2006; Corballis, 2015; Nenert et al., 2017; Zhang et al., 2018). These qualities make the left hemisphere well-suited to handling complex, co-ordinated motor and cognitive functions, like those involved in interpreting, planning, and articulating speech. This is because their predisposition towards sequential and segmental analysis increases the likelihood of left-hemispheric structures recruiting functions like speech sound representation, sound-meaning mapping, rapid auditory processing, syntactic processing, reading, and handling abstract concepts (Lindell, 2006; Poremba, 2006; Sussman, 2015; Riès et al., 2016).

However, the left (or dominant) hemisphere does not exclusively house all language-related functions. Some structures in the right (or non-dominant) hemisphere are also instrumental in supporting language (Lindell, 2006). These include the supplementary motor cortex, pars opercularis, supramarginal gyrus, posterior and medial superior temporal gyrus, precuneus, orbitofrontal cortex, posterior cingulate, frontal eye field, and parts of the prefrontal cortex (Bottini et al., 1994; Harpaz et al., 2009; Riès et al., 2016; Patel et al., 2018; Zhang et al., 2018). Some of these right-hemispheric regions contribute to oro-motor movement and co-ordination, and others are primarily involved in interpreting figurative language, supporting semantic processing, handling concrete concepts, and integrating emotional-affective and pragmatic information contained in prosody (volume, pitch, speech rate and rhythm) and paralinguistic cues, (Salmelin & Sams, 2002; Lindell, 2006; Harpaz et al., 2009; Sussman, 2015; Patel et al., 2018). The right posterior superior temporal gyrus (the structure in the non-dominant hemisphere corresponding to or *homologous to* Wernicke's area), for instance, is typically associated with ambiguity resolution when we encounter less frequent meanings of homonymous words e.g., 'bank' as in 'riverbank' vs. 'bank' as in 'bank account' (Harpaz et al., 2009; Riès et al., 2016). The right hemisphere's predisposition towards global, holistic processing may explain why it plays a greater role in language perception and comprehension than in production. Interpreting a pre-existing incoming linguistic signal using a range of auditory and visuospatial cues may be achievable by less specialised, more holistic processing than retrieving, constructing, and articulating an original sequential utterance (Lindell, 2006; Riès et al., 2016; Corballis, 2017).

There is a recognised association between atypical or weak lateralisation for language and neurodivergence, cognitive impairment, and mental illness (Nicholls et al., 2005; Bishop, 2013; Corballis, 2013; Lindell, 2016, 2020). However, the relationship between lateralisation and cognitive ability is not entirely straightforward. One theory suggests that the degree to which language-related functions are lateralised is primarily determined by the typicality or atypicality of an individual’s neurodevelopmental experiences, with genetic influences playing a less influential role (Bishop, 2013, p. 10). This position aims to account for why some but not all individuals with weakly lateralised language processing show cognitive impairments. However, the very high three-way coincidence of weak or atypical lateralisation of language functions amongst people with comorbid neurodivergencies and/or mental health conditions suggests that there may be a common genetic (endophenotypic) cause in some people (Lindell, 2020). It is possible that experience and genetics may interact differently amongst neurotypical and neurodivergent populations and therefore influence lateralisation in different ways or to different degrees.

There is some evidence that language processing becomes increasingly bilateral in advanced age (see Nenert et al., 2017). However, this effect may be restricted to the temporal and parietal regions of some (but not all) right-handed men, since this effect was only observed within this group in Nenert et al.'s (2017) research, but not in women or in left-handed men. An MRI (Magnetic Resonance Imaging) study involving 224 right- and left-handed men and women aged 18-76 found robust population-level patterns of activity in partially overlapping, left lateralised task-dependent networks during semantic decision-making and verb generation (Nenert et al., 2017). These networks involve inferior portions of the left frontal cortex (including the inferior frontal gyrus), the left angular and cingulate gyri, both superior temporal gyri, and portions of the cerebellum.

It may be that certain language-related functions like rapid auditory processing and word retrieval, are more sensitive to the direction of lateralisation than others (Lindell, 2016; Riès et al., 2016; Cantiani et al., 2019). A higher proportion of 6-month-olds at risk of language impairment later in childhood show right lateralised neural activity during rapid auditory processing – a vital skill for interpreting and acquiring spoken language – compared with their typically developing peers (Cantiani et al., 2019). The laterality, amplitude, magnitude, and timing of babies’ neural responses to rapid, sequential auditory input are strong predictors of receptive and productive vocabulary size at 20 months (Cantiani et al., 2019). Unilateral lesions (resulting from illness or brain injury) in areas involved in word retrieval are associated with more significant language impairment than lesions affecting word selection, and this may be because word selection can be achieved through less specialised and more bilaterally represented types of cognition than can word retrieval (Riès et al., 2016). Rapid auditory processing and word *retrieval* may be more specialised operations, requiring co-operation and co-ordination of several generalised capacities.

Where damage occurs in brain areas supporting less flexible or generalised language-related functions, language outcomes may be mildly to severely compromised (Raja Beharelle et al., 2010; Lindell, 2016; Riès et al., 2016). Foetal exposure to alcohol can cause a range of cognitive impairments requiring continuous support, resulting from irreparable damage and atypicalities in the asymmetry, morphology, cellular composition, and functionality of areas associated with language processing, particularly in the temporo-parietal regions (Lindell, 2016). Foetal alcohol exposure and foetal or early life brain injury can trigger neighbouring or homologous areas of the brain to recruit or support functions typically performed by the damaged area/s, though these functions are typically performed less proficiently, perhaps because the neurons in these areas are over-loaded or are imperfectly equipped for the types of structural analysis required (Johnson, 2000; Mareschal et al., 2007; Morillon et al., 2010; Raja Beharelle et al., 2010; Lindell, 2016; Riès et al., 2016). Where typically left-hemispheric functions are re-recruited to right hemisphere homologues, this can result in more compromised performance (Raja Beharelle et al., 2010). Interestingly when a typically left-hemispheric function is retained by a left-hemispheric area or comes to be represented bilaterally, language outcomes may be significantly less disrupted regardless of the extent or severity of the left-hemispheric lesion (Raja Beharelle et al., 2010). Where brain injury occurs later in life, language processing may be more severely compromised still, and individuals affected may experience fluent (Wernicke's) or dysfluent (Broca's) aphasia or dysphasia (Nishitani et al., 2005; Raja Beharelle et al., 2010; Szaflarski et al., 2014). Throughout the lifespan, neuroplasticity declines, at first more rapidly and then more gradually, meaning that recovery from brain injury becomes increasingly reliant on the brain's ability to repair already-lateralised existing neural architecture structures and connections, rather than forming new connections and recruiting new or homologous areas (Szaflarski et al., 2014; Olulade et al., 2020).

In cases where a less specialised function that can be performed equally proficiently by homologous areas in either hemisphere (e.g., pitch perception) shows a less frequently observed pattern of lateralisation, language outcomes may be entirely undisrupted (Johnson, 2000; Morillon et al., 2010; Minagawa-Kawai et al., 2011; Riès et al., 2016). Some language-related capacities are sufficiently general that they can be recruited to one hemisphere or the other according to the precise function that they perform (Minagawa-Kawai et al., 2011). Amongst babies acquiring tonal languages, near-infrared spectroscopy (NIRS) has shown that pitch variation in speech elicits greater left-hemispheric activity, but greater right-hemispheric activity in speakers of non-tonal languages (Minagawa-Kawai et al., 2011). In tonal languages, pitch carries linguistic information about semantic contrasts whilst in non-tonal languages pitch carries more paralinguistic emotional state information. Sensory perception of pitch relies on a relatively simple and generalised ability of the auditory cortex to detect changes in frequency (Sussman, 2015). It is possible that attaching significance to pitch may be a secondary and more specialised ability building on this sensory percept (Tierney & Nelson, 2009) and involving interactions between the auditory cortex and a network of either predominantly right-hemispheric areas involved in emotional processing or predominantly left-hemispheric areas involved in sound-meaning mapping. Pitch processing may then become lateralised and localised through interactions between individual experience, the genetically determined proclivities of neurons within each hemisphere for handling certain types of sensory information, and the type of processing required for the individual to accurately interpret the meaning carried by variations in pitch (i.e., global/holistic vs. sequential/segmental) (Minagawa-Kawai et al., 2011).

These findings together highlight the importance of diversity of sampling in laterality research. In the past, there has been a precedent in research in social and psychological disciplines for over-sampling neurotypical, healthy, white, monolingual English-speaking, mid-SES, and right-handed males aged 18-40 for reasons of consistency, which can limit the generalisability of resulting findings (e.g., see Nenert et al., 2017). Sampling more representatively in laterality research will help us to build a more comprehensive picture of laterality and lateralisation and allow us to identify variability and developmental change. Since weak or atypical lateralisation is often associated with cognitive and linguistic impairments, identifying individual- and group-level differences in laterality may help us to identify individuals requiring additional support (Holowka & Petitto, 2002a; Bishop, 2013; Lindell, 2016, 2020). Understanding when and how lateralisation proceeds and how it contributes to typical development could inform how we construct therapies and interventions and when and how we implement them. If lateralisation of linguistic and cognitive functions is not fixed at birth, then there is scope to implement interventions that encourage more typical lateralisation, for example, by using tasks designed to elicit activity in brain areas associated with speech sound processing, social communication, and constructing and interpreting complex motor activities (e.g., the superior temporal gyrus and inferior frontal gyrus). Stimulating these areas could result in strengthening of the parts of the brain that are typically major players in supportin linguistic and cognitive functioning (Van Der Knaap & Valk, 1990; Mareschal et al., 2007; Petanjek et al., 2011; Williamson & Lyons, 2018).

Many of the brain areas that support language processing are modality-independent, having the flexibility to accommodate language-related functions in the auditory-vocal modality, as in spoken languages, and in the visuo-motor modality, as in signed languages, reading, and writing (Inubushi & Sakai, 2013; Newman et al., 2015). Functional Transcranial Doppler sonography (fTCD) has revealed similar degrees of left lateralised neural activity during language production in hearing-impaired signing children and typically hearing children who use spoken language, matched for language ability (Payne et al., 2019). Functional Magnetic Resonance Imaging (fMRI) has identified a common left lateralised network involved in processing signed and spoken language, comprising the inferior frontal gyrus, the lateral premotor cortex, and the superior temporal sulcus (Meister et al., 2007; Inubushi & Sakai, 2013; Newman et al., 2015). Positron Emission Tomography (PET) has identified a common left lateralised network for semantic processing during speech perception and reading, which includes the anterior fusiform gyrus, the anterior superior temporal sulcus, the lateral temporal pole, and the posterior portion of the junction between the temporal, parietal, and occipital lobes (Spitsyna et al., 2006). That aspects of language processing are modality-independent may help us to understand how and why the brain is able to retain certain functions like communicating by written word after brain injuries that impair speech processing.

Brain areas involved in language processing are also domain-general, being involved in multiple different functions. For example, the left inferior frontal gyrus is involved in representing and producing sequential manual movements as well as sequential vocalisations (Nishitani et al., 2005; Hodgson et al., 2021). Areas in the left superior temporal gyrus are involved in musical cognition and generalised auditory processing as well as speech sound representation (Keenan et al., 2001; Griffiths & Warren, 2002; Meyer et al., 2012). Some of these brain areas – including the classical language areas – have homologues in at least some extinct hominins including Homo erectus, Homo habilis, and Homo naledi, and some living great apes and Old World monkeys, including Japanese and rhesus macaques, baboons, chimpanzees, and gorillas (e.g., see Petrides & Pandya, 2009; Wu and colleagues 2011a, 2011b). Anatomical and neuroimaging research has mapped the morphology of non-human primates' brains. Endocasts showing the external morphology of extinct hominin's brains have been created by using digital modelling or plaster casting to map the internal morphology of the brain casing, Hominin and non-human primate homologues of classical language areas are associated with similar functions to those areas in modern humans, and exhibit similar leftward asymmetries in size, morphology, and connectivity (Holloway, 1983; Gannon et al., 1998; Broadfield et al., 2001; Cantalupo & Hopkins, 2001; Griffiths & Warren, 2002; Jürgens, 2003; Poremba et al., 2003; Nishitani et al., 2005; Wu et al., 2006; Taglialatela et al., 2008; Petrides & Pandya, 2009; Wu et al., 2011a; Wu & Pan, 2011b; Corballis, 2015; Holloway et al., 2018; Becker et al., 2021; Becker et al., 2022). Some authors have characterised left lateralisation of language as a sudden mutation or 'speciation event' in modern human evolution (Crow, 2008) but comparative research has provided abundant evidence of significant continuity across species and across evolutionary time.

This research is particularly important: the knowledge that language is supported by modality-independent and domain-general capacities contributes to our understanding of how language first emerged in our ancestors and how it emerges in our children by making use of pre-existing neural architecture capable of performing multiple roles when participating in multiple distinct networks. This knowledge can also help us to understand how best to support people facing socio-cognitive and linguistic difficulties resulting from developmental impairment or brain injury. Language may not be a single unique behaviour underlain by a discrete language-specific brain-based capacity (c.f. Petitto and colleagues, 1999, 2022a, 2002b, 2004) but rather a unique system in which modified forms of ancient social and survival-related behaviours and capacities are coupled together in new and special ways. It may be from this myriad of complex and highly specific *interactions between* ancient domain-general capacities – more than the specific capacities themselves – that human language arises (Thelen & Smith, 1994; Serrien et al., 2006; Bunton, 2008; Häberling et al., 2016; Nenert et al., 2017).

The remainder of this chapter will review evidence from research concerning the laterality of vocal and communicative behaviours in humans and non-human animals and consider whether the balance of this evidence most closely aligns with the hypothesis of phonological development put forward in this thesis or with the other competing hypotheses discussed in Chapter 2. Laterality research has used a range of behavioural, neuroimaging, neurosurgical, and neuroanatomical techniques. The present study concerns orofacial asymmetry analysis, and so the following sections will first review previous orofacial asymmetry research and then comment on how the findings of this research relate to those of research using alternative techniques. Methods for analysing facial asymmetry are also reviewed.

Laterality of language-related functions: adult research

Studies examining asymmetry in lip openings during speech and facial expressions in adults have found measurable behavioural evidence of left lateralisation for meaningful and meaningless speech and right lateralisation for emotional facial expressions (Graves et al., 1982; Wyler et al., 1987; Wylie & Goodale, 1988; Graves & Landis, 1990).

Orofacial asymmetry research

Orofacial asymmetry analysis works on the principle that each hemisphere of the brain is primarily responsible for controlling the muscles on the contralateral side of the lower face and body (Rinn, 1984). A number of different methods have been developed for examining orofacial asymmetry in photographs, video data, and still-frames extracted from video data.

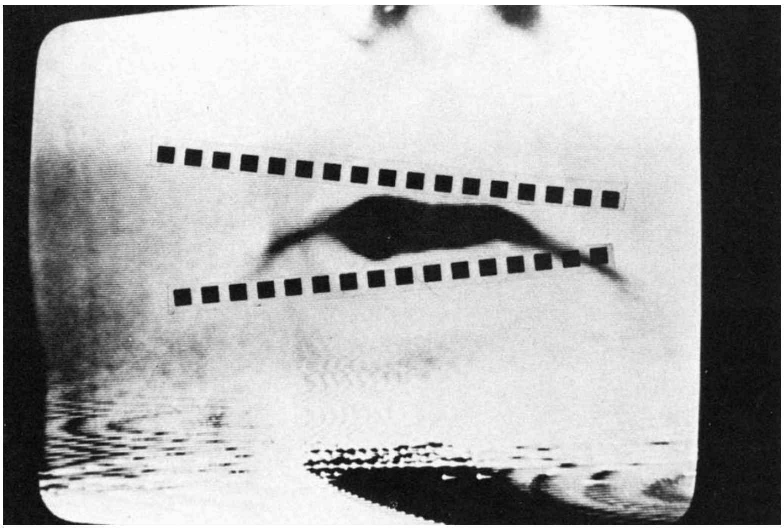


Figure 2 Image showing the method developed by Graves et al. (1982) for analysing asymmetry in lip openings via the angle created between two rulers, one placed parallel to each lip. This image is reproduced from Graves et al. (1982, p. 373).

Graves et al. (1982) developed a method for analysing orofacial asymmetry (asymmetry in lip openings) wherein two rulers are placed onto an enlarged still-frame of a person's mouth during a vocalisation, one aligned with each lip, with the angle between them indicating whether the left or right side of the mouth (or neither) has opened more widely (see Figure 2). Still-frames are then coded -1 if the left side opened more widely, +1 if the right side opened more widely, or 0 if each side of the mouth appeared to open equally widely). Laterality Indices – values indicating which side of the mouth opened more widely – are then calculated for each participant by summing the number of instances of greater right-sided (R) and left-sided (L) mouth opening and the number of instances of equal opening on each side of the mouth (E) and inserting these values into the formula: (R – L) / (R + L + E). Positive laterality indices indicate left hemisphere dominance, whiles negative laterality indices indicate right hemisphere dominance. Overall, 76% of 150 right- and left-handed adults tested by Graves et al. showed greater right-sided mouth opening, indicating left lateralisation of capacities supporting speech production. In a visual image description task, 85% of men and 67% of women showed left lateralisation. In alliterative word generation task, 86% of participants showed left lateralisation, with no significant sex difference. The sex difference in the image description task may have resulted from experiment design: the image description task elicited more emotional responses, particularly amongst women, some of whom displayed amusement (Graves et al., 1982). In a separate task, researchers elicited one spontaneous smile from each participant, and these showed right lateralisation (Graves et al., 1982)[16]. This was the foundational study for the laterality research with babbling babies carried out by Holowka and Petitto (2002), and the findings of Graves et al.'s experiments shaped Holowka and Petitto's interpretation of their own findings.

Wyler et al. (1987) sought to confirm Graves et al.'s (1982) assertion regarding the relationship between language and emotion using the same manual method for analysing orofacial asymmetry. Forty-six right-handers (23 female) participated in three emotional or visuospatial tasks involving describing visual memories (e.g., remembering a first date) or imagined visuospatial scenarios (e.g., how an individual would hug a person or what a clock looks like) and three purely linguistic tasks including explaining an abstract conceptual distinction (the word 'knowledge' vs. the word 'intelligence') and retrieving phonologically related word from memory (words beginning with /m/). Consistent with Graves et al. (1982), purely linguistic tasks showed stronger left lateralisation whilst visuospatial and emotional language tasks showed weaker evidence of left lateralisation, suggesting greater involvement of the right hemisphere in these latter tasks.

---

[16] A body of research discussed in more depth in Chapter 6 has examined the laterality of emotion and the balance of the evidence from this research confirms that processing associated with emotion and emotional facial expressions takes place in a network of predominantly right hemispheric structures (see Fox & Davidson, 1988; Wylie & Goodale, 1988 amongst others; Best & Queen, 1989; Nagy, 2012).

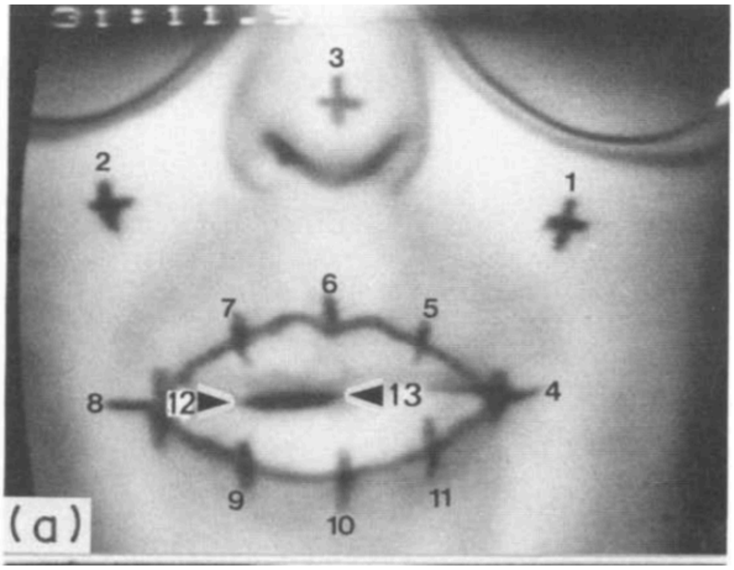


Figure 3 Image showing the method for analysing orofacial asymmetry via lip displacement used by Wolf and Goodale (1987) and Wylie and Goodale (1988). This image is reproduced from Wolf and Goodale (1987, p. 379).

Another method involves measuring lip displacement away from markers placed on the nose tip and cheeks over a sequence of consecutive or non-consecutive still-frames (Wolf & Goodale, 1987; Wylie & Goodale, 1988). This method is illustrated in Figure 3. A study using this method examined lip openings during sequential and non-sequential, ‘linguistic’ and non-linguistic oral gestures and made complementary findings (Wolf & Goodale, 1987). Unlike in Graves et al.’s (1982) study, utterances in the ‘linguistic’ condition in this experiment were the meaningless syllables ‘ba’, ‘ma’ and ‘pi’ rather than natural language or isolated words. The non-linguistic gestures were the commands ‘open’ (the lips), ‘blow’, and ‘close’ (the lips. The right side of the mouth was found to begin to open first and to open more widely than the left during syllable production and non-speech oral gestures containing similar mouth movements (Wolf & Goodale, 1987). This effect was slightly, though non-significantly, more pronounced in isolated speech than in isolated non-speech. The strength of the right mouth bias *was* significantly increased, however, during utterances containing a string of syllables compared with isolated syllables. Sequences, requiring more planning and articulatory effort, involved more strongly left lateralised processing (Wolf & Goodale, 1987). It would be interesting to know whether babies also show similarly increased lateralisation for reduplicated and variegated babble as compared with singleton babble, since these subtypes of babble are taken to present a weightier cognitive and articulatory load and to indicate the emergence of more sophisticated phonological skills. Unlike in Graves et al.'s (1982) experiments in which men showed stronger left lateralisation, stronger left lateralisation was observed in women than in men in Wolf and Goodale's (1987) study.

A third technique, used by Hausmann et al. (1998) involves dividing an image of the mouth into two *hemimouths* and measuring the height at the highest point of the lip opening in each hemimouth[17]. This method has found evidence of robust left lateralisation across men and women when producing isolated mono- and polysyllabic words and non-words regardless of emotional valence (Hausmann et al., 1998). In continuous or repetitive speech, significant left lateralisation was found only in men (Hausmann et al., 1998).

Taken together, these findings suggest that the act of representing, planning, and producing syllables may be left lateralised independently of any other types of cognition associated with meaningful language use (the syntactic, semantic, pragmatics, or symbolic). There may be a parallel here between the meaningless, reference-less speech acts used in some of these studies and babies' babble. If the auditory-articulatory component of speech is left lateralised independently of emotional cognition or other types of linguistic cognition then this may call into question whether left lateralisation constitutes sufficient evidence to support the position that babble is "fundamentally linguistic" (Holowka & Petitto, 2002a, p. 1515). Instead, this finding may suggest that language comes to be left lateralised because – or perhaps *if* – the domain-general capacity for representing, planning, and producing sequential, vocalised mouth movements is recruited to the left hemisphere at an earlier stage in development. This capacity and the parts of the brain that support it may then represent a relatively low-level but nonetheless critical cognitive-motoric function, which, once established, may serve as a sort of foundation or as a sub-system in the development of higher-level functions that make use of it like meaningful speech (Thelen & Smith, 1994; Bates, 1999; Tierney & Nelson, 2009).

---

[17] Hausmann et al. (1998) do not provide figure illustrating this technique for measuring orofacial asymmetry.

Sex differences

It is possible that findings of less strongly expressed laterality or more bilateral language processing amongst women may reflect genuine sex differences in cerebral organisation and processing, but neuroimaging and stroke research have not found unequivocal evidence to support this claim (Frost et al., 1999; Clements et al., 2006; Wallentin, 2018). Alternatively, apparent sex differences may be related to hormonal fluctuations. Cyclical increases in the amounts of progesterone and oestradiol in the body during the luteal phase of the menstrual cycle may be associated with weaker lateralisation of language-related functions, whilst subsequent drops in hormone levels during menstrual phase may promote differential strengthening of the lateralisation of language-related functions (Fernández et al., 2003; Lindell, 2006). This may explain why laterality research variously finds weaker or stronger lateralisation for language in women than in men (Wolf & Goodale, 1987; Hausmann et al., 1998; Frost et al., 1999; Sommer et al., 2004; Clements et al., 2006; Lindell, 2006; Kaiser et al., 2007; Sommer et al., 2008).

Other techniques

Since the time of these studies, their key findings have been confirmed and extended by behavioural, neuroimaging, and neuroanatomical research.

Laterality and handedness

Further evidence has been found suggesting that left lateralisation of language-related functions is largely independent from handedness. In fTCD research, some 90-96% of right-handers, 73-78% of left-handers, and 89% of ambidextrous adults show left lateralised activity during word generation (Knecht et al., 2000). Whilst the proportion of individuals showing atypical laterality for language processing is said by the authors to increase linearly as a function of handedness, the overall tendency towards left lateralisation remains robust. Extensive behavioural testing has demonstrated that only around 20% of left-handers show right lateralised language processing (Van Der Haegen & Brysbaert, 2018). Whilst left-handers may show more intraindividual variability in the laterality of different language-related functions, the population-level tendency towards left lateralisation persists even where left-handers are deliberately over-sampled (Nenert et al., 2017; Van Der Haegen & Brysbaert, 2018). fMRI has found that only 4.7-12.6% of right-handers and only 23.4-25.9% left handers show atypically lateralised activity in the frontal lobes, inferior frontal gyri, and temporo-parietal areas during semantic processing and verb retrieval (Nenert et al., 2017). However, there remains a strong precedent in laterality research for associating species-level left laterality for language with our tendency towards right-handedness. The strength and nature of this relationship therefore merits full consideration in a thesis concerning laterality and, as such, will be discussed further detail later in this chapter.

Meaningfulness and meaninglessness: words and syllables

Magnetoencephalography (MEG) has shown that isolated or serial production of meaningful bisyllabic words, silent vowel gestures, and single-movement silent non-speech mouth movements alike are associated with left lateralised activation in the face area of the motor cortex (Salmelin & Sams, 2002). However, word production generates more focal activation. This may be because bisyllabic words involve more sequential processing. A later MEG study found left lateralised activity supporting isolated and serially produced CVCV-structured words and non-words and single-movement silent non-speech mouth movements, with no effect of category, though the leftward asymmetry strengthened as a function of utterance duration and complexity (Saarinen et al., 2005). Here, there is a clearer parallel between these meaningless, reference-less CVCV utterances and babies' babble.

While both are left lateralised, meaningful and meaningless sequential mouth movements may be supported by distinct task-dependent networks (Jessen, 1999; Bunton, 2008, Kern et al., 2019). Subdural electrocorticography[18] (ECoG), fMRI and PET have identified distinct and dispersed networks of neurons involved in producing speech, smiling, laughing, and moving the tongue or lips (Bunton, 2008; Kern et al., 2019). Producing meaningless mouth movements involves more bilateral activation while meaningful speech and emotional expressions show stronger lateral biases (Bunton, 2008). With regard to perception, exposure to meaningful but not meaningless speech elicits activity in the left posterior

---

[18] ECoG involves taking electroencephalographic (EEG) readings using electrodes inserted directly into the exposed cortex during surgery.

cingulate gyrus, the left angular gyrus, and the bilateral anterior and medial superior temporal gyri and precuneus (Jessen et al., 1999). Speech and language therapies using non-speech mouth movements and vocalisations are typically less effective than techniques employing naturalistic speech and stimulating socio-cognitive development (Bunton, 2008; Watson & Lof, 2008; Wilson et al., 2008; Kent, 2015; Kern et al., 2019; Parra-López et al., 2022).

Taken together, these findings lend further support to the idea that planning and articulating sequential movements of the mouth may be recruited to the left hemisphere quite independently of other types of linguistic cognition like processing meaning. That distinct task-dependent neuronal networks participate in meaningful speech processing but not in other oro-motor behaviours in adults may also pose a problem for theories that language evolved or develops out of feeding behaviours or manual gestures (MacNeilage & Davis, 1993; Corballis, 2003). It would be informative to know whether the network/s supporting the *production* of meaningless syllable production are more similar to those supporting meaningful speech production or meaningless speech perception. In either case though, the fact that jaw co-ordination in vocal behaviour becomes adult-like between 12-30 months, many months earlier than jaw co-ordination in feeding (Wilson et al., 2008) suggests that feeding is supported by a different network or networks to both meaningless and meaningful sequential vocalisation. By 12 months, babies have typically already accumulated much experience of producing meaningless sequential vocalisations, and by 30 months, will have begun to represent a growing range of meaningful sequential vocalisations as well (Oller, 1980; Stark, 1980; Vihman 2014).

Syllables and speech sounds

Other neuroimaging research has further explored at what level/s left lateralised sequential processing of meaningful and meaningless speech is performed.

Simultaneous fMRI and EEG has shown that frequencies of electrical activity in the brain associated with syllabic processing (delta-theta oscillations – ‘brainwaves’) and phonemic processing (gamma oscillations) are expressed more strongly in the left auditory, somatosensory, articulatory motor, and inferior parietal areas than in their right hemisphere homologues when the brain is at rest (Morillon et al., 2010). Delta-theta oscillations increase in the left lip and hand motor cortical areas when we are exposed to natural language, but gamma oscillations do not, suggesting that fluent, meaningful language processing may involve articulatory parsing at the level of syllables but not at the level of phonemes (Morillon et al., 2010). Morillon et al. suggest that syllabic processing may be "hard-wired" (pp. 18688, 18691) and that this may explain why even hearing impaired babies, with little to no extrinsic experience of speech, are motivated to babble (Petitto et al., 2004, cited by Morillon et al., 2010). By contrast, the authors suggest that phonemic processing may begin only once babies have accumulated enough feedback from their own vocalisations and enough ambient language exposure to recognise matches between the two and to begin to create categorical speech sound representations. However, simultaneous fMRI and EEG with premature neonates has found evidence of left lateralised activation in the inferior frontal and posterior temporal regions in response to consonantal place of articulation contrasts in isolated syllables at 28-32 weeks gestational age, before gestational cortical development is complete (Mahmoudzadeh et al., 2013).

So then, we may possess some innate sensitivity to the acoustic properties of both syllables and speech sounds. Mahmoudzadeh et al. (2013) do note, though, that the patterns of neural activity seen in these neonates are not adult-like, pointing to a significant role for articulatory and perceptual experience during development in hemispheric and regional specialisation. It is possible that the amount of phonetic information contained in an utterance may also influence which neural mechanisms are employed to parse it, with the longer sequences of natural language promoting syllabic parsing and the shorter sequences of isolated, meaningless syllables allowing both phonemic and syllabic parsing.

fMRI has identified distinct but overlapping areas of the posterior superior temporal gyrus that support speech sound production and perception in adults (Buchsbaum et al., 2001). The posterior superior temporal gyrus represents categorically only those acoustic differences in the speech signal that are contrastive in the ambient language/s rather than the full continuum of all possible acoustic differences (Chang et al., 2010), suggesting that precisely how the brain represents and analyses speech sounds is influenced by individual experience and is not wholly innately prespecified. Intracranial electroencephalography (iEEG) has identified clustering patterns in the neurons involved in processing specific speech sounds according to manner of articulation (primarily) and to place of articulation (secondarily) (Mesgarani et al., 2014). This primary clustering pattern may result from motoric similarities and/or order of acquisition. Stops are easier to articulate reliably, requiring complete closure, and are typically the earliest consonants acquired so may be represented in the brain first. Fricatives require more precise articulation and are typically acquired later. Representing fricatives alongside stops may stimulate the networks supporting speech sound representation to expand and differentiate, forming new synapses or recruiting new neurons in order to maintain distinctions between representations for speech sounds with different neuromuscular action plans. The secondary clustering pattern may emerge later, perhaps around the time when babies develop Vocal Motor Schemes as these post-date the earliest consonants and centre around a preferred place of articulation (McCune & Vihman, 1987). The posterior superior temporal gyrus possesses connections to the inferior frontal gyrus and the parietal operculum (Buchsbaum et al., 2001; Hickok & Poeppel, 2007; Corballis, 2015). These projections are strongly left lateralised and facilitate dynamic integration of acoustic and articulatory information (Hickok & Poeppel, 2007). The neighbouring anterior and medial portions of the superior temporal gyrus are more heavily involved in sound-meaning mapping and are more bilaterally organised (Hickok and Poeppel, 2017). These findings, too, support the suggestion that purely phonetic and phonological processing is lateralised independently from other language-related functions.

Summary

Language-related functions are predominantly left lateralised in adults, but this tendency towards left laterality is not absolute, with a minority of people showing right lateralised or bilaterally represented language processing. Nor is the laterality of language-related functions wholly innate or unchanging. Rather, laterality may weaken with age (in right-handed men); may fluctuate with hormonal changes (in women); and may continue to develop some time after birth through interactions between genetics and experience. Behavioural and neuroimaging research have provided evidence that different language-related functions are supported by distinct task-dependent networks of neurons. Each of these networks may be lateralised to a greater or lesser degree than other networks with which it is connected and coactivated during language processing.

The parts of the brain that are involved in these task-dependent networks are largely domain-general. For example, the inferior frontal gyrus supports visuomotor and auditory processing and working memory functioning in interpreting, representing, retrieving, and co-ordinating complex, sequential, linguistic and non-linguistic oral and manual movements (Müller & Basho, 2004; Nishitani et al., 2005; Hodgson et al., 2021). The left posterior superior temporal gyrus is involved in musical cognition and generalised acoustic processing as well as in perceiving and representing the sounds of speech (Binder et al., 1996; Griffiths & Warren, 2002; Elmer et al., 2016; Binder, 2017). The left superior temporal sulcus is involved in different types of domain-specific and domain-general social cognition, auditory processing, face processing, and theory of mind as well as in acquiring the phonological inventory of the native language (Deen et al., 2015; Newman et al., 2015). The posterior portion of the left cingulate gyrus, typically associated with attention, emotional regulation, and online adaptation to intrinsic and environmental changes, is also active when parsing meaningful speech (Jessen et al., 1999; Leech et al., 2012; Leech & Sharp, 2014). Critically, none of these areas are capable of performing language-related functions in isolation. Rather, it is the extensive connectivity and the highly specific interactions between these areas that afford humans the ability to process language.

Babble provides evidence to suggest that a left lateralised network for *sequential vocalisation* develops during the first year of life (Holowka & Petitto, 2002a). However, since babies exhibit little-to-no evidence of possessing the other capacities involved in language processing before 12 months of age, this left laterality in and of itself may not provide sufficient evidence to suggest that babble is underlain by a dispersed, left lateralised network of neuronal clusters supporting an innate capacity for processing language.

The following sections will explore the findings of some developmental and comparative laterality research that may offer some insight into how domain general brain areas have come to support functions involved in language processing.

Laterality of communicative functions in non-human animals

Species-level cerebral and functional asymmetries are not unique to humans (Corballis, 2008). Lateral asymmetries can be seen in some animal behaviours that may be construed as evolutionary precursors to language (e.g., Cantalupo & Hopkins, 2001; Marie et al., 2018; Becker et al., 2021, 2022). Production of communicative vocalisations is left lateralised in birds of the *Passeridae* family and frogs, and perception of species-specific vocalisations is left lateralised in mice, rhesus macaques, Japanese macaques, and chimpanzees (Vallortigara et al., 1999; Corballis, 2015). The findings of laterality research with non-human animals are of great value and significance to studies concerning language ontogeny and phylogeny and should be given due consideration in any attempt to formulate theories of language development and evolution. That a leftward lateral bias for processing vocal communicative signals of varying degrees of complexity pre-dates the evolution of language and modern humans by several million years (Corballis, 2008) adds further weight to the argument that left laterality in babble may not be an indicator of an innate and species-specific preparedness for language. This section will focus primarily on evidence from non-human primates.

Orofacial asymmetry research

In our close cousins, chimpanzees, orofacial asymmetry research has revealed evidence of differential neural processing underlying artificial, learned vocalisations, as compared with their species-typical calls (Losin et al., 2008; Wallez et al., 2012). Like in humans, the lower half of the face in non-human primates is contralaterally controlled (Hook-Costigan & Rogers, 1998; Wallez & Vauclair, 2012; Fernandez-Carriba et al., 2002a, 2002b). In pant-hoots and food barks, produced by both wild and captive chimpanzees, the left side of the mouth opens more widely than the right, indicating greater involvement of the right hemisphere (Losin et al., 2008; Wallez et al., 2012). By contrast, vocalisations taught to captive chimpanzees by their human carers for cross-species communication, including extended grunts, ‘raspberries’, and ‘kiss’-like sounds, are produced with greater right-sided mouth openings, indicating left hemisphere dominance (Losin et al., 2008; Wallez et al., 2012). These attention-getting sounds are learned through a sort of cultural transmission between humans and chimpanzees and are produced referentially and in a goal-directed way so were taken in these studies as analogues in chimpanzee behaviour for language in humans. Chimpanzees’ species-typical vocalisations are widely held to convey emotional state information and have, until quite recently, been considered to be largely involuntary, so were used as analogues for spontaneous emotional facial expressions in humans.

Losin et al. (2008) and Wallez et al. (2012) suggest that learned, referential, communicative vocalisations and emotional processing were likely lateralised differently in our common ancestor with chimpanzees (from whom we diverged 6.1-6.7 MYA[19], timetree.org). Neither Losin et al.'s (2008) nor Wallez et al.’s (2012) studies involved chimpanzees under the age of 5, nor do the authors comment on when these chimpanzees learned these attention-getting sounds or how proficient they were at producing them (though it may be presumed that they were proficient enough for them to be identifiable and categorisable). This means that their research does not address whether the differential hemispheric specialisation that they observed is in evidence from birth (and so possibly under biological control) or whether this pattern emerges over time (i.e., via experience-dependent processes of development). More recently, species-typical chimpanzee calls, including pant-hoots and grunts, have been found to show evidence of vocal learning, referentiality, and of intentional on-line modification to convey different shades of meaning to different types of intended audience, and chimpanzees have been observed to extract social and environmental information from the acoustic properties of conspecifics’ calls and modify their own behaviour accordingly (Slocombe & Zuberbühler, 2007; Slocombe et al., 2010b; Schel et al., 2013a; Watson et al., 2015; Fedurek et al., 2015b). It would be interesting for future research to examine whether the degree of lateralisation seen in species-typical and -atypical vocalisations varies as a function of learning or development.

Losin et al. (2008) and Wallez et al. (2012) used an adapted version of an alternative method for analysing orofacial asymmetry. This method was first developed by Hook-Costigan and Rogers (1998) for use in laterality research on emotion in marmosets and has also been used to analyse laterality in species-typical communicative and emotional vocalisations in chimpanzees, baboons, and macaques (Fernández-Carriba et al., 2002a; 2002b; Wallez & Vauclair, 2012). Being designed for use with naturalistic audio-visual data, several exclusion and adjustment criteria are used to identify still-frames that are suitable for analysis (see Chapter 4). The total area of the lip opening is then traced using image editing software

---

[19] (million years ago)

(most commonly, Adobe Photoshop) and the mouth is divided into two hemimouths using facial landmarks like the inner eye corners (see Figure 4). The area of each hemimouth is then found by counting the pixels in each hemimouth and an Oral Asymmetry Index is calculated for each image.

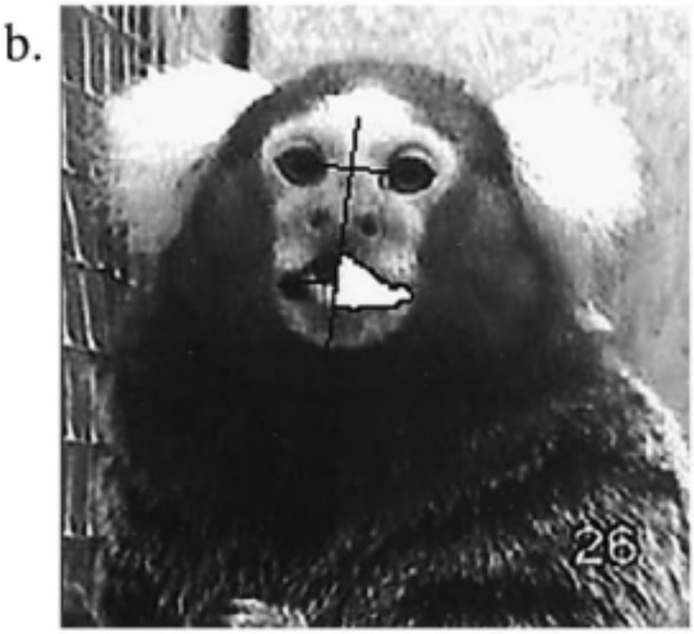


Figure 4 Image illustrating how the inner eye corners are used to identify the midline of the face enabling measurement of the left hemimouth. This image reproduced from Hook-Costigan and Rogers (1998, p. 1267).

Naturalistic observation allows researchers to collect more spontaneous, ecologically valid laterality data than experimental paradigms, but also presents certain complications since research participants are free to move and interact with their environment and other. This means that the face may sometimes be angled away from the camera, may move during a vocalisation, or may be partially obscured by some object like food, the hands, toys, or tree branches. This is the method adapted for use with human infants in the present study. The present study favours this method for its rigour and its applicability across age groups and species, and with both naturalistic and experimental data, which gives the advantage that resulting findings may be directly compared.

Other techniques

Laterality research using a range of other techniques has provided some support for Losin et al. (2008) and Wallez et al.'s (2012) findings. Similarities in the size, morphology, asymmetry, and function of the inferior frontal gyri and plana temporalia in humans and their homologues in other extant great apes (observed in anatomical and neuroimaging research) and extinct hominins (inferred from skull endocasts) are consistent with the claim that some left-hemispheric specialisation for learned, intentionally communicative vocalisations pre-date modern humans and language (Gannon et al., 1998; Jürgens, 1998; Cantalupo & Hopkins, 2001; Hopkins & Cantalupo, 2003; Taglialatela et al., 2008). In fact, functional and cellular similarities within less morphologically human-like homologues of these areas in Old World monkeys, taken along with the evidence of vocal learning in these monkeys, suggest that some hemispheric and regional specialisation for learned communicative vocal and manual behaviours dates back further, to a common catarrhine ancestor some 27.0-30.6 MYA (Kumar er al., 2022; Masataka & Fujita, 1989; Marshall et al., 1999; Crockford et al., 2004; Russell et al., 2013; Fischer et al., 2015; Watson et al., 2015; Fischer et al., 2020; Vernes et al., 2021).

Behavioural research

Adult rhesus macaques show left lateralised head orientation responses to the calls of conspecifics, but right lateralised responses to calls of birds with which they are not familiar (Hauser & Andersson, 1994)[20]. Infant rhesus macaques show less specialisation: 54.1% exhibit left lateralised responses to conspecifics' calls but 60% exhibit left lateralised responses to unfamiliar bird calls. Laterality in macaques, like humans, is not wholly pre-determined before birth but involves some experience-dependent process of development during infancy and childhood (Hauser & Andersson, 1994). Grey mouse lemurs (one of the earliest evolving extant primates) also show some evidence of left lateralisation for perception of conspecifics' calls but only amongst males and only for negatively valenced

---

[20] It is acknowledged that in 2007 M. D. Hauser was found guilty of eight counts of academic misconduct and his work has been subject to multiple retractions. According to information made publicly available at https://www.federalregister.gov/documents/2012/09/06/2012-21992/findings-of-research-misconduct, the study cited here was not amongst those found to contain fabricated data. The findings of this study are discussed here because laterality research before and since 2007 has cited this study and because similar laterality studies have made similar findings. However, these findings should be considered with caution.

calls, with positively valenced calls and the calls of other animals eliciting right lateralised responses in both sexes (Scheumann & Zimmermann, 2008). Females in this study showed no clear population-level left lateralisation for conspecifics' calls, though it is possible that this may reflect hormonal fluctuations in female grey mouse lemurs of the type discussed by Lindell (2006) in relation to humans. A closely related species – the lesser mouse lemur – has an adaptable[21] 33-63 day oestrus cycle, with a >25 day luteal phase, during which progesterone levels are raised (Perret, 1986). Increased progesterone is associated with weakened lateralisation in humans (Lindell, 2006).

So then, left lateralisation for processing species typical calls was likely present in a common catarrhine ancestor some 27.0-30.6 MYA (timetree.org) and possibly, but not necessarily, in a common strepsirrhine[22]-haplorrhine[23] ancestor, around 71.4-77.5 MYA (timetree.org). However, research with other non-human mammals suggests that this trait may pre-date even a common primate ancestor, dating as far back as 89.6-97.4 MYA (timetree.org). Preferential listening paradigms have been used in conjunction with heart monitors to gauge which hemisphere is most strongly activated when domesticated animals hear emotionally valenced calls of conspecifics and heterospecfics. The animal participant is fitted with a heart monitor and provided with a bowl of food as a central fixator. Positive, negative, and neutral vocalisations of conspecifics and/or heterospecfics are played from speakers positioned to the left and the right of the animal whilst they eat. Vocalisations are then played from the speakers positioned to the right and left of the animal. Since auditory perception is contralateral in these species (Siniscalchi et al., 2008; 2016, 2018), elevated heart rate and greater interest when a sound is played from the lefthand speaker indicates right lateralised processing of the auditory signal and vice versa. Domesticated cats show strongly left lateralised responses to positive or neutral cat calls, but strongly right lateralised responses to dog barks expressing negative emotion (Siniscalchi et al., 2016). Negative cat calls and neutral dog calls elicit weakly right lateralised responses. In domesticated dogs, exposure to other dogs' barks elicits left lateralised responses while sounds from nature elicit right lateralised responses (Siniscalchi et al., 2008). Negatively

---

[21] The lesser mouse lemur oestrus cycle decreases to c.30 days in length when in isolation but may increase up to more than 60 when in close physical proximity to other females.

[22] Our common ancestor with lemurs

[23] Our common ancestor with gibbons, tarsiers, Old World monkeys, and New World monkeys

valenced barks show a weaker effect suggestive of greater right-hemispheric emotional processing (Siniscalchi et al., 2016). Demonstrating once again the role of experience-dependent processes in lateralisation, negatively valenced *human* vocalisations elicit stronger right lateralised cardiac and behavioural responses in domesticated dogs, whilst positively valenced or neutral human vocalisations elicit strongly left lateralised responses (Siniscalchi et al., 2018).

So, laterality research with non-human animals may demonstrate that a right-hemispheric specialisation for emotional processing and a left-hemispheric specialisation for processing information contained in meaningful communicative vocalisation is a particularly ancient trait. Some of these findings may provide support for Valence theory – the theory that the right hemisphere processes only negative emotion, while positive emotion is processed in the left hemisphere (see Sackeim et al., 1982). Alternatively, it may be that hemispheric dominance in these examples is determined by the kind of processing necessary and/or sufficient for the listener to respond appropriately (Hook-Costigan & Rogers, 1998). For negatively valenced vocalisations, global processing may provide more useful information e.g., an aggressive warning call may prompt the listener to retreat to a more mutually comfortable distance regardless of the species of either interlocutor. For neutral or positively valenced vocalisations, closer sequential analysis of the constituent parts of a vocalisation may be more beneficial. This may be largely restricted to conspecifics' calls, since these may contain a wider range of more personally relevant meanings for the intended listener, but exceptions may be seen in heterospecifics who spend a good deal of time together. For example, dogs can be observed to extract salient, meaningful chunks of information from continuous speech e.g., "Puck can you sit down please? Good dog. Have a biscuit.". Findings of left lateralised responses for species-typical calls in non-human mammals may seem to conflict with Losin et al. (2008) and Wallez et al.'s (2012) findings. However, the species-typical vocalisation examined in these chimpanzee studies were selected precisely because they were strongly emotionally valenced, with pant-hoots being aggressive warnings and food-barks being expressions of excitement. Whether emotions are processed differentially according to valence in humans (i.e., *Valence Theory*) is controversial but, as in the case of these chimpanzees, the balance of the evidence would seem to suggest not. Valence theory and its relevance to the present thesis are considered further in Chapter 6.

Neuroanatomical research

As well as behavioural similarities, similarities in the anatomy of the frontal and temporal cortices have been identified between humans and other animals from whom we diverged up to at least 100 MYA (see Cantalupo & Hopkins, 2001; Marie et al., 2018; Becker and colleagues, 2021, 2022; Kumar et al., 2022). The brains of grey seals (divergence at 89.6-97.4 MYA) show similar sulcal and gyral morphology, cortical asymmetry, and cortical size relative to body size to human brains (Hoeksema et al., 2021). A gene that is heavily implicated in acquiring linguistic, vocal, and motor skills in humans – namely, FOXP2 – is expressed the grey seal cortex as well as the human cortex and has been implicated in vocal learning in this seals (Hoeksema et al., 2021). The brains of chimpanzees, bonobos, orangutans, and gorillas, and certain Old World monkeys show leftward asymmetries in the size and organisation of the planum temporale and inferior frontal gyrus, although to a lesser degree than humans (Gannon et al., 1998; Cantalupo & Hopkins, 2001; Marie et al., 2018; Neubauer et al., 2020; Becker et al., 2021; Becker et al., 2022). For brevity, the remainder of this section will explore similarities between the classical language processing areas in humans and their homologues in extant and extinct primates.

The planum temporale

The left planum temporale, housed in the superior temporal gyrus, is enlarged relative to its right hemisphere homologue in humans and non-human primates (Cantalupo & Hopkins, 2001). In humans, the left planum temporale has historically been considered a language-specific brain area (Gannon, 1998; Harpaz et al., 2009; Binder, 2017). However, it may serve a more domain-general role or roles in speech comprehension via the combined efforts of the task-dependent networks for auditory, articulatory, and semantic processing in which it participates (Thelen & Smith, 1994; Griffiths & Warren, 2002; Tierney & Nelson, 2009).

In humans, an especially enlarged left planum temporale at the individual level can be associated with musical proficiency and absolute pitch (Gannon et al., 1998; Keenan et al., 2001; Meyer et al., 2012). Musical cognition is classically associated with the right hemisphere but being trained to analyse technical features of music may give rise atypical lateralisation (Glanville et al., 1977; Koelsch et al., 2005; Ono et al., 2011). Enhanced individual-level asymmetries in the planum temporale may provide some insight into precisely how a species-level leftward asymmetry in this region contributes to language development and evolution. Expert musicians with an especially enlarged left planum temporale show a particular aptitude over untrained individuals in language-related tasks involving phonetic processing (Meyer et al., 2012). Expert musicians may develop more differentiated representations of a wider range of meaningful sound categories, affording them a heightened sensitivity to acoustic distinctions not salient to the untrained ear and brain (Buchsbaum et al., 2001; Chang et al., 2010; Mesgarani et al., 2014; Weiss et al., 2018).

Similarly, primates with more differentiated communicative behaviours are observed to show greater cerebral asymmetries than primates with more limited ranges of calls (Cantalupo & Hopkins, 2001). Ninety-four percent of chimpanzees show similar leftward asymmetries to humans in planum temporale size and gyral and sulcal organisation, suggesting that we have inherited a heavy reliance on sequential acoustic processing and an accompanying, highly developed neural substrate for doing it from our common ancestor c.8 MYA (Gannon et al., 1998).

It was previously suggested that this trait did not date back to a common catarrhine ancestor since vervet monkeys, rhesus macaques, and bonnet macaques show no evidence of population-level asymmetries (Cantalupo & Hopkins, 2001). However, another species of Old World monkey – the Olive baboon – *does* show leftward asymmetries in the planum temporale that are markedly similar to those in humans, which may date this trait to between 30-40 MYA (Marie et al., 2018). Leftward asymmetries in size and density of grey vs. white matter are seen in similar proportions of newborn and adult baboons and humans, and this asymmetry increases in degree during post-natal development (Becker et al., 2021; Becker et al., 2022). Baboons make extensive use of a wider range of intentional vocalisations and manual gestures in their social behaviour than some other Old World monkeys and so must be able to interpret and construct meaning in these gestures and vocalisations (Arnold & Zuberbühler, 2006; Ouattara et al., 2009; Marie et al., 2018). It is possible that other Old World monkeys with less highly differentiated vocal behaviour may be less reliant on the categorical representational abilities of the left planum temporale specifically. For example, in rhesus macaques, hearing conspecifics' calls elicits significant activation in a network of areas including the whole superior temporal gyrus, not restricted to the planum temporale, along with areas within the parietal and prefrontal cortices and the limbic area (Poremba et al., 2003). Certain areas in this network are also involved in visuospatial processing in macaques (Hecht & Parr, 2015). This latter finding illustrates how specialised capacities that support vocal communicative behaviour may arise out of interactions between domain-general and modality-independent capacities even in primates often considered to possess much 'simpler' brains than humans (Thelen & Smith, 1994; Poremba et al., 2003; Tierney & Nelson, 2009).

The inferior frontal gyrus

The left inferior frontal gyrus is enlarged in living great apes (Cantalupo & Hopkins, 2001; Corballis, 2015) and was also enlarged in certain pre-modern hominins (e.g. see Broadfield et al., 2001). Earlier hominins lived in increasingly complex social groups, likely maintained by increasingly sophisticated systems for vocal and manual communication (Broadfield et al., 2001; Holloway et al., 2018; Dusseldorp & Lombard, 2021). Leftward asymmetries in the size and morphology of the planum temporale and inferior frontal gyrus reflecting those seen in adult humans are seen in endocasts of Homo erectus skulls spanning almost a million years. These include *Hexian Homo erectus*, dated to 412 KYA[24] (Wu et al., 2006); *Sambungmacan 3 Homo erectus*, dated to 100 KYA – 1MYA[25][26] (million years) (Broadfield et al., 2001; Delson et al., 2001); *Nanjing 1 Homo erectus*, dated to 0.58–0.62 MYA (Wu et al., 2011a); and *Zhoukoudian Homo erectus*, dating from 0.4-0.8 MYA (Wu & Pan, 2011b). Homo habilis endocasts dating from 1.8-2.0 MYA exhibit the earliest evidence of an enlarged left inferior frontal gyrus in the direct lineage of modern humans (Holloway, 1983; Tobias,

---

[24] KYA denoting thousand years ago

[25] MYA denoting million years ago

[26] Sambungmacan 3 Homo erectus is difficult to date owing to having been found in the banks of an Indonesian river (the Solo) in 1977 and illegally transported from Indonesia to America and back in the 1990's (Delson et al., 2001; Márquez et al., 2001). Exposure to water compromises the accuracy of techniques used for carbon-dating and the conditions under which the skull was transported (humidity, temperature, packaging etc.) are not known.

1987). Working on the principal that evolutionary change takes place gradually over generations via descent with modification, it seems likely that some gradual complexification and differentiation of function was underway in much earlier ancestors. Indeed, whilst showing a smaller brain volume than other hominins, *Homo naledi* skulls dating to 236-335 KYA show similar morphology and organisation of the inferior frontal gyri and lateral orbital gyri to modern humans (Holloway et al., 2018). *Homo naledi* are not direct ancestors of modern humans but co-existed with us for c.64 million years, having branched off 900 KA – 2.0 MYA (Holloway et al., 2018).

The reliability digital or physical endocasts created from fossilised skulls is sometimes debated, relying, as it does, on small samples of skull specimens that are often damaged or incomplete (Beaudet, 2017). However, similarities in the size, morphology, and function of the inferior frontal gyri in our living evolutionary cousins adds weight to the findings of fossil research.

Great apes have more complex social group structures than Old World monkeys, and this is reflected in their frontal cortices: in bonobos, chimpanzees, and gorillas, the left pars opercularis is similarly enlarged and alike in sulcal and gyral morphology to that in humans (Cantalupo & Hopkins, 2001). In chimpanzees, the left inferior frontal gyrus is more enlarged relative to the right than in Old World monkeys and exhibits more differentiated connectivity to the temporal and parietal regions compared its homologue in macaques – area F5 (Hecht & Parr, 2015). The human inferior frontal gyrus shows more differentiation still (Hecht & Parr, 2015).

In at least four species of primate, homologues of the inferior frontal gyrus contain visuomotor and auditory mirror neurons, and these are implicated in learning, planning, and producing complex, sequential, communicative vocalisations and gestures, and other oral and manual movements (Vihman, 2002; Crockford et al., 2004; Rizzolatti & Craighero, 2004; Watson et al., 2015). Infant Japanese and rhesus macaques exhibit a period of vocal learning lasting several months, during which they gradually modify their own calls towards adult targets (Masataka & Fujita, 1989). Unlike macaques, the more differentiated mirror systems of chimpanzees and humans allow us to represent intransitive actions (Hecht & Parr, 2015).

Understanding indirect causation is vital for acquiring language but also for engaging in other aspects of great ape social and cultural life like adhering to group customs for social bonding or learning how to craft and use specialised tools. Chimpanzees acquire species-typical communicative, social, and survival-related behaviours through observation and interaction and will voluntarily and intentionally tailor their communicative vocalisations and gestures to specific audiences in order to convey different meanings or shades of meaning and elicit different types of response from their conspecific listener/s (Marshall et al., 1999; Leavens, 2003; Crockford et al., 2004; Slocombe & Zuberbühler, 2007; Slocombe et al., 2010b; Russell et al., 2013; Schel et al., 2013a; Schel et al., 2013b; Fischer et al., 2015; Watson et al., 2015; Fedurek et al., 2015a; Fedurek et al., 2015b). Infant and adult chimpanzees can also acquire arbitrary non-species-typical vocalisations and gestures for communicating and co-operating with heterospecifics (Losin et al., 2008; Wallez et al., 2012). Captive chimpanzees will spontaneously selectively modify the rate and frequency of their vocalisations on-line to attract or redirect humans' attention or to request food, and may selectively use gesture only in joint attention scenarios (Hostetter et al., 2001). Wild chimpanzees also use learned communicative vocalisations to promote social cohesion (Hotstetter et al., 2001).

Lip-smacking, vocal learning, and proto-grammar

One particular communicative vocalisation type in primates – lip-smacking – has been the focus of some considerable research and may offer some insights into language ontogeny and phylogeny (Morrill et al., 2012; Ghazanfar & Takahashi, 2014; Fedurek et al., 2015; Takahashi et al., 2015;). Lip-smacking, when produced during the initiation of a grooming episode, is associated with longer and more frequently reciprocated grooming episodes compared with episodes during which chimpanzees produce fewer or no lip-smacks (Fedurek et al., 2015a). Furthermore, lip-smacks are typically produced only when chimpanzees are interacting face-to-face i.e., once joint attention is established (Fedurek et al., 2015a). Language and the neural activity that supports it may have emerged through gradual complexification and differentiation of behaviours like lip-smacking. This may have come about as an adaptive response to social challenges associated with living in especially large and complex social groups which needed to work cohesively for survival, and in which each individual plays one or more social roles in relation to others around them (Snowdon, 2017).

In macaques, lip-smacking is produced with co-ordinated phonation and lip movements, occurring at a rhythm of 3-8Hz, similar to human speech (Ghazanfar & Takahashi, 2014; Takahashi et al., 2015). Infant macaques exhibit a period of vocal learning that has been liked to human phonological development[27]: their early lip-smacks are unco-ordinated, inconsistent, and produced at a slow rate, but become more consistently adult-like with time and articulatory practice (Morrill et al., 2012). Interestingly, vocal and feeding behaviours exhibit different rhythmic properties and developmental trajectories in infant macaques, suggestive of distinct neural substrates (Morrill et al., 2012). It would be interesting to know whether a similar process of vocal learning takes place for lip-smacking in infant chimpanzees and whether the laterality of lip-smacking strengthens or becomes more consistent with time and articulatory practice in either species.

Another interesting finding is that at least two[28] species of primate more distantly related to humans exhibit some evidence of proto-semantic and proto-syntactic sensitivity (Ouattara et al., 2009; Zuberbühler, 2020). Campbell's monkey has six distinct categories of call, which are produced in isolation or concatenated according to five distinct patterns to convey detailed context-specific information about threats to survival (Ouattara et al., 2009). These threat warnings include information about whether the threat is predatory, non-predatory, or environmental; where it is located; whether the species of a predator or non-predator is known to the caller; and what the species is (Ouattara et al., 2009). Putty-nosed monkeys, one of the smallest Old World monkeys, have only two distinct categories of call associated with two predators (leopards and crowned eagles). These can be reduplicated or concatenated according to some loosely defined principles to communicate what and where a predator is and whether the group should attempt escape or remain still to avoid detection (Arnold & Zuberbühler, 2006). Interpreting and constructing meaning in these kinds of concatenated vocalisations involves representing and retrieving associations

---

[27] Morrill et al. (2012) compare macaques' early lip-smacks to babble but note that babble *is* regular and rhythmic, so a closer comparison may be marginal babble.

[28] Work finding evidence of proto-grammatical ability in cotton-top tamarins has been carried out by Hauser and colleagues but is not discussed here because of its uncertain relationship to the work involved in Hauser's (2007) misconduct trial. https://www.federalregister.gov/documents/2012/09/06/2012-21992/findings-of-research-misconduct

between discrete vocal sounds and events or objects in the world, and implicitly recognising that the nature of the transitions between these meaningful 'chunks' are themselves also meaningful (Ouattara et al., 2009; Zuberbühler, 2020). This suggests that in Old World monkeys with very limited vocal repertoires, area F5 is sufficiently sensitive and differentiated to carry out these highly complex kinds of computation in the absence of any behaviour that we would call linguistic (Zuberbühler, 2020).

Language-related functions in the inferior frontal gyrus may have evolved through gradual differentiation and complexification of more generalised functions in our early ancestors (Petrides & Pandya, 2009). The human right inferior frontal gyrus and the bilateral macaque area F5 support retrieval of visuospatial information from memory (Poremba et al., 2003; Müller & Basho, 2004; Nishitani et al., 2005; Petrides & Pandya, 2009; Sussman, 2015; Häberling & Corballis, 2016; Häberling et al., 2016; Hodgson et al., 2021). Properties of neurons in the primate left hemisphere may make the left pars triangularis (BA 45) more efficient than the right at retrieving more complex, hierarchically structured auditory and visuospatial information from memory (Petrides & Pandya, 2009). The left pars opercularis (BA 44) and its macaque homologue may have evolved complementarily to the pars triangularis, to be adept at "transforming" this remembered information into a motor acts and "instantiating" it (Petrides & Pandya, 2009, p. 13). In adult chimpanzees, producing communicative manual gestures and vocalisations – like lip-smacks – is associated with increased activity throughout the left inferior frontal gyrus including BA 44 and 45 but also the left pars orbitalis (BA 47/12) and frontal pole (Taglialatela et al., 2008). In humans, the human pars orbitalis is involved in phonological, semantic and syntactic processing (De Carli et al., 2007) and the frontal pole is involved in social cognition, memory, attention, and selecting appropriate and beneficial responses to environmental stimuli (Kovach et al., 2012; Bludau et al., 2014).

Summary

Laterality research with non-human animals has helped to redefine the relationship between language and structures in the left hemisphere that support it. The left inferior frontal gyrus and superior temporal gyrus began to expand and differentiate quite early in primate evolution to accommodate representations of a greater variety of sequential gestures of the hands and mouth, and to recognise more nuanced relationships between acoustic information and objects, phenomena, and other organisms in the world (Griffiths & Warren, 2002; Nishitani et al., 2005; Luef, 2018). However, leftward functional and anatomical asymmetries may not be a direct index of complexity of behaviour; instead, a stronger predictor of complexity may be the degree of individual-level variability in laterality within a species (Neubauer et al., 2020). While some cerebral asymmetries are apparent from birth, the tendency for the calls of conspecifics but not heterospecifics to be processed by the left hemisphere may emerge through experience-dependent processes rather than being purely genetically determined. Species-typical vocalisations may contain more information promoting survival and social cohesion and may therefore need to be deconstructed and interpreted in a way that heterospecifics' vocalisations usually do not need to be (with notable exceptions like dogs and humans).

The research reviewed in this section of the chapter favours the Old Parts, New Machine hypothesis. Meaningful communicative behaviours in non-human animals show left lateralisation in the absence of language, further weakening the argument that left lateralisation in babies' early meaningless syllables is an indicator of inherent linguistic-ness. That other primates including Old World monkeys exhibit evidence of distinct neural substrates for vocal and ingestive behaviours also weakens the hypothesis that language evolved or develops out of jaw movements involved in feeding. Instead, the research reviewed here points to an evolutionary and developmental scenario wherein the complex cognitive and motoric capacities involved in language emerge out of highly specific interactions between highly differentiated but domain-general parts of the human brain involved in social cognition and vocalisation. A fundamental characteristic of dynamic biological systems is that the way they behave arises out of the interactions between the subsystems that participate in them (Thelen & Smith, 1994). These interactions are influenced and constrained by the properties of the participating subsystems and the experiences and needs of the organism.

Laterality in phonological development and language acquisition: research with babies and children

Laterality research involving babies and children offers insights into the ontogeny and sometimes also the phylogeny of language lateralisation (e.g., Dehaene-Lambertz et al., 2002, 2010; Bogolepova & Malofeeva, 2001; Weiss et al., 2018; Wang et al., 2021). Some structures and functions whose roles will later adapt to include supporting language may to be left lateralised from or before the time of birth, suggesting a role for genetic determination (Mareschal et al., 2007). Other functions may be recruited by each hemisphere through experience-dependent processes of behavioural development and associated cerebral specialisation (Bishop, 2013).

Orofacial asymmetry, innateness, and a brain-based language capacity

Despite its suitability for naturalistic laterality research with babies and toddlers, little orofacial asymmetry analysis has been carried out with these groups. One important orofacial asymmetry study by Holowka and Petitto (2002a) found evidence of differentiated cerebral specialisation for babble and emotional facial expressions in pre-linguistic babies under the age of 1. French and English babies (n=5 in each group) exhibited group- and individual-level left lateralisation during reduplicated babble, but right lateralisation during smiles, and no clear lateralisation during other vocalisation types not meeting the criteria for babble, termed *non-babble*. Holowka and Petitto's (2002a) study and its findings are of particular import to the present study because the present study builds on Holowka and Petitto's (2002) work. The authors used the method developed by Graves et al. (1982) that was outlined earlier in this chapter, and they draw comparisons between Graves et al.'s findings and their own. Holowka and Petitto interpret the finding that babble is left hemisphere dominant as support for their hypothesis that babble is ”fundamentally linguistic” (p. 1515).

Other techniques

Some neuroimaging and neuroanatomical evidence partially align with Holowka and Petitto's (2022a) findings and their interpretation thereof. fTCD (functional Transcranial

Doppler sonography) has found left lateralisation amongst 3-10-year-olds when asked to describe an animation (Hodgson et al., 2016). NIRS has identified left lateralised activity in areas in the temporal and frontal cortices in 3-30-year-olds for semantically-related word retrieval, with significantly more bilateral activity observed in adolescents and adults compared with the youngest children sampled (Paquette et al., 2015). These authors have interpreted their findings as evidence that laterality is genetically determined, and that language processing is left lateralised from the time of birth. Evidence from behavioural paradigms and EEG has provided some evidence to support this conclusion (Hahn, 1987). For example, at 3 months, speech syllables elicit left lateralised cardiac responses, whilst musical notes elicit right lateralised cardiac responses (Glanville et al., 1977). Two-month-olds show left lateralised activity in the temporal lobe when exposed to speech, and right lateralised activity in the temporal lobe when exposed to music (Dehaene-Lambertz et al., 2010). The brains of newborns exhibit leftward asymmetries and minor individual differences in sulcal and gyral organisation reflecting those seen in adult human and non-human primates. For example, the left planum temporale and left frontal and temporal cortices of newborns are enlarged relative to the right and contain a higher density of neurons (Witelson & Pallie, 1973; Chi et al., 1977; Wada, 1977; Bogolepova & Malofeeva, 2001). In babies aged <28 days, the pars opercularis and pars triangularis (BA 44 and 45) are well-defined in both hemispheres with the left homologues containing a higher density of neurons than the right, and BA 45 showing a higher density still than BA 44 (Bogolepova & Malofeeva, 2001). Taken together, the findings that some parts of the brain show leftward asymmetries at the time of birth could be interpreted as evidence that the human brain is somehow biologically specialised to acquire language (see Witelson & Pallie, 1973; Petitto and colleagues, 1991, 2002a, 2004). (However, it is argued in this thesis that this interpretation downplays the role of *in utero* sensory experience in neural development.)

Asymmetries in the temporal cortex may even pre-date birth (Chi et al., 1977; Bogolepova & Malofeeva, 2001; Herin-Hanit et al., 2001; Corballis et al., 2013). In ultrasound scans, the foetal left hemisphere is larger than the right by around 20-22 weeks' gestation (Hering-Hanit et al., 2001). Comparison of foetal brains at 10-44 weeks' gestation shows that leftward asymmetries in the planum temporale and transverse temporal gyri (or *Heschl's gyri*, involved in audition) emerge during the third trimester from around 31 weeks' gestation (Chi et al., 1977). There is even some evidence of brain asymmetry during the first trimester. For example, the choroid plexus shows leftward asymmetries in size and morphology in foetal brains at 11-13 weeks' gestation (Corballis, 2013). The choroid plexus is a network of blood vessels in the ventricles of the brain whose function is to produce cerebrospinal fluid, which nourishes and cleanses the brain and provides shock absorption; to keep the cerebrospinal fluid from combining with the blood supply to the brain; and to secrete substances that trigger growth (Javed et al., 2021; Telano & Baker, 2021). During the first trimester, the choroid plexus grows large relative to body size and declines in relative size with age (Corballis, 2013). Corballis (2013) suggests that its early asymmetry and its role in growth may indicate that it is involved in generating the leftward asymmetries seen in the temporal region by the third trimester (Chi et al., 1977). By this reasoning, it may also exert some influence over the asymmetry and differences in neuron density seen in the inferior frontal gyrus in newborns (Bogolepova & Malofeeva, 2001)

Early-life leftward asymmetries in neuroanatomy and language-related functions have been interpreted by some authors as an indication that humans possess some innate left lateralised linguistic ability or predisposition (Witelson & Paille, 1973). Witelson and Pallie (1973, p. 646) write that leftward asymmetry of the planum temporale in newborns "indicates that the infant is born with a pre-programmed biological capacity to process speech sounds". The authors argue that this asymmetry must be genetically determined, claiming that it manifests in the absence of any type of formative experience. Petitto and colleagues have suggested that left lateralisation in babble indicates that there exists a "brain-based language capacity" that is triggered by exposure to language (Petitto & Marentette, 1991, p.1495; Petitto et al., 2004, p.46; see also Holowka & Petitto, 2002a), arguing that motoric accounts of babble emergence fail to explain the similarity and continuity between babble and language. Paquette et al. (2015, p. 117) write that language lateralisation is determined by "structural and functional properties of the human brain with little reorganisation occurring in development".

An alternative perspective

If it is true that laterality is genetically determined, and that language processing is left lateralised from the time of birth, then this would lend significant support to the Linguistic hypothesis. However, evidence from other developmental research runs counter to these claims.

Pre-natal experience-dependent cerebral specialisation

Cerebral specialisation occurring before birth does not preclude the possibility of influence from experience as Paquette et al. (2015) suggest (see Dehaene-Lambertz et al., 2010; Minagawa-Kawai et al., 2011). Audition and short-term auditory memory become functional between 20-25 weeks' gestation and foetuses can be observed to respond to changes in acoustic stimuli in the womb from 28 weeks' gestation (Querleu et al., 1988; Previc, 1991; Graven & Browne, 2008). By 37 weeks' gestation, foetuses show differential responses to frequently experienced familiar speech sounds and infrequently-experienced sounds (DeCasper et al., 1994). In 0-3-day-olds, low-pass filtered ambient language speech elicits stronger bilateral temporal activity than non-ambient language speech (May et al., 2011). In 1-day-olds, exposure to unfiltered ambient language speech elicits left lateralised activity in the temporal region, whilst hearing a non-native language elicits greater right temporal activation, and non-speech sounds elicit more bilateral responses (Vannasing et al., 2016). Differences between Vannasing et al. and May et al.'s findings may result from the stimuli used: Vannasing et al.'s (2016) stimuli contained sequential, segmental information, which is processed more efficiently by the left hemisphere, whilst May et al.'s (2011) stimuli were filtered to retain prosodic information, which is commonly associated with the right hemisphere though can be processed in either (see Minagawa-Kawai et al., 2011), but relatively little segmental information. In either case, differential neural responses to ambient language vs. non-ambient language speech strongly suggests a role for pre-natal experience in processes involved in lateralisation as well as genetics.

Foetal and placental position may also influence cerebral specialisation (Previc, 1991). Noting that approximately 66% of foetuses rest with the right side facing the front of the uterus, Previc (1991) suggests that the right ear may receive higher quality auditory input and the right arm may have a greater possible range of movement. This may mean that the left hemisphere is a) more frequently activated and b) provided with more variable sensory information than the right. Repeated activation strengthens neural pathways and increases their likelihood of reactivation in future (Van Der Knaap & Valk, 1990; Hayakawa et al., 1991; Mareschal et al., 2007; Su et al., 2008; Williamson & Lyons, 2018; Grotheer et al., 2022). Variability in the input builds more differentiated and localised neuronal networks for processing incoming information (Paterson et al., 2006; Mareschal et al., 2007). Previc suggests that this trend in foetal position may account for the species-level tendencies towards left laterality for language and right handedness.

Further evidence that foetal movement is involved in differential hemispheric specialisation may come from a foetal orofacial asymmetry study by Reissland and colleagues (2014). Between 24-36 weeks' gestation, bilateral foetal mouth movements decrease by 11% each week, being superseded by more asymmetric mouth movements (Reissland et al., 2014). The left cerebral hemisphere becomes larger and more differentiated than the right around this time, so one might predict that more mouth movements with greater right-sided lip openings should be seen, However, the opposite effect is observed. Reissland et al. suggest that the increase in activity on the left side of the mouth may reflect developments in lateralisation for emotional expressivity as the foetus responds to their internal emotional state with movements of the face. Interestingly, some lateralisation for emotional facial expression continues well beyond gestation and throughout the second year of life (Schuetze & Reid, 2005): during this time, babies and toddlers show increasingly right lateralised negatively valenced emotional facial expressions. Unlike Holowka and Petitto's (2002), Schuetze and Reid found no clear lateralisation for positively valenced emotional facial expressions. This discrepancy may have resulted from the different methods used to measure orofacial asymmetry in these studies. Holowka and Petitto measured the angles created by the opening of the lips (see also Graves et al., 1982; Wyler et al., 1987) while Schuetze and Reid divided the mouth into two hemimouths and measured the relative area of each hemimouth (see also Hook-Costigan & Rogers, 1998; Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). This latter method – measuring hemimouth area – is more sensitive to asymmetries not easily visible to the naked eye and has been described by Fernandez-Carriba and colleagues as more objective than the former method.

In order for experience-dependent development to take place, the developing organism must actively interact with its environment and respond to sensory input and feedback (Thelen & Smith, 1994; Mareschal et al., 2007; Tierney & Nelson, 2009). Whilst genetics may exert more influence over cerebral organisation than experience during gestation, given the rapid proliferative brain development that occurs during gestation and the fairly limited range of sensory information available to the foetus (Tierney & Nelson, 2009), foetal arm, face, and general body movements may constitute sufficiently agentive interactions between the foetus and their environment to instigate some degree of cerebral specialisation. Perceptual and productive capacities are supported by largely the same multimodal networks of brain areas and it has been suggested that perceptual asymmetries resulting from interactions between information processing biases and early experiences may give rise to later productive asymmetries (Dehaene-Lambertz et al., 2002; Vihman, 2002; Dehaene-Lambertz et al., 2010; Morillon et al., 2010; Minagawa-Kawai et al., 2011; Sussman, 2015). The availability of certain perceptual capacities during gestation – notably, audition – may mean that perceptual lateralisation during gestation can influence and constrain lateralisation for productive behaviours including babble and speech many months post-partum.

Post-natal experience-dependent cerebral specialisation

Counter to the claim that laterality is wholly determined from the time of birth, there is much evidence of progressive lateralisation of language-related functions post-natally as well as pre-natally (Ramsay, 1980, 1984; Best & Queen, 1989; Rothbart et al., 1989; Mundy et al., 2000; Schuetze & Reid, 2005; Workman et al., 2006; Dehaene-Lambertz et al., 2010; Nagy, 2012; Rosselli et al., 2014; Emerson et al., 2016; Hodgson et al., 2016; Lindell et al., 2017; Weiss et al., 2018; Reynolds et al., 2019; Yang et al., 2019; Olulade et al., 2020; Wang et al., 2021). Orofacial asymmetry research has even found some evidence of measurable changes in the laterality of the processes support the processing and production of facial expressions – which emerge long before babble and language – persisting beyond the first year (Best & Queen, 1989; Rothbart et al., 1989; Schuetze & Reid, 2005; Workman et al., 2006; Nagy, 2012; Lindell et al., 2017).

Babies and toddlers

Significant structural changes take place in the brain during the first year, at first in the temporal and parietal cortices, with especially salient changes seen in the frontal cortices during the latter half of the first year (Bogolepova & Malofeeva, 2001; Paterson et al., 2006; Mareschal et al., 2007). Whilst somewhat adult-like leftward asymmetries in the size, morphology, and activity of neural structures (particularly in the temporal and inferior frontal cortices) are seen in the newborn brain, *connectivity* between areas and hemispheres is less evident and less established in infancy than in adulthood (Dehaene-Lambertz et al., 2002; Paterson et al., 2006; Mareschal et al., 2007; Perani et al., 2011). Additionally, the newborn brain exhibits some morphological asymmetries that are not adult-like in the left occipital and temporo-parietal regions (Witelson & Pallie, 1973; Chi et al., 1977; Wada, 1977).

Between <28 days, 6 months, and 1-2 years, heterochronic and non-linear changes in relative neuronal density and cortical thickness take place in the pars opercularis and pars triangularis in both hemispheres (Bogolepova & Malofeeva, 2001). During the first year, the size, cortical thickness, and neuronal density of the right homologues overtakes that of the left homologues, which may reflect the fact that babies' phonological experiences during the first 6 months are largely receptive (Bogolopova & Malofeeva), before the emergence of sequential vocalisation or an idiosyncratic phonological system. This may mean that vocal sounds are processed more globally and are supported by more distributed and/or less consistent neural activity (Bogolepova & Malofeeva, 2001; Mareschal et al., 2007; Tierney & Nelson, 2009). Between 6-24 months, the left pars triangularis (BA 45, on the left inferior frontal gyrus) develops more rapidly than the pars opercularis (BA 44, on the left inferior frontal gyrus) in either hemisphere, and this may reflect developments in articulatory planning ability and the rapid proliferation in productive phonological and speech-related behaviours typical of this period (Bogolepova & Malofeeva, 2001).

When exposed to forward and backward speech, sleeping and wakeful 3-month-olds show left lateralised activity in the superior temporal sulcus, superior temporal gyrus, planum temporale, and temporal pole (Dehaene-Lambertz et al., 2002). Wakeful babies also show left lateralised activity in the left planum temporale and right prefrontal cortex (Dehaene-Lambertz et al., 2002). Adults show different patterns of activity in these areas when exposed to forward speech whilst awake, but no comparable activation when sleeping or when exposed to backward speech (Dehaene-Lambertz et al., 2002), indicating that some functional and perhaps structural reorganisation takes place between infancy and adulthood. Morillon et al. (2010) claim that lateralisation for syllables processing may be innate. Dehaene-Lambertz et al.'s findings do not speak to ideas of innateness. Whilst suggesting that syllable processing may be *lateralised* by or before 3 months, they also indicate that some process of experience-dependent regional cerebral specialisation for processing syllables – likely involved effort and attention – is yet to take place between 3 months and adulthood. Ambient language exposure, and maternal speech specifically, begins to influence lateralisation during gestation (Glanville et al., 1977) and continues to do so after birth. In 2-month-olds, only maternal speech elicits greater activation in the posterior portion of the temporal lobe, as well as in right-hemispheric emotional processing areas compared with listening to speech in an unfamiliar voice (Dehaene-Lambertz et al., 2010). Sensory processing during sleep may represent one experience-dependent neural process or mechanism that is instrumental to phonological development. EEG research has demonstrated that newborns can acquire non-ambient-language vowel boundaries when exposed to novel categorical stimuli whilst sleeping, and that this categorical discrimination is robust against changes in voice or pitch (Cheour et al., 2002). Adults do not share this capacity for processing speech sounds when sleeping and the authors suggest that this capacity may decline over the first year of life. This capacity may become less and less useful over time as babies come to spend longer periods of time awake and attentive, which may also have the consequence that the brain requires more concentrated rest when sleeping.

Some regional and hemispheric changes in the structures supporting language persist into the period when babies begin to acquire words, mediated by individual productive and perceptual experience. For example, NIRS has revealed greater activation in the right posterior superior temporal sulcus than in its left homologue in 11-month-old babies exposed to sound-symbolic stimuli that contains labial sounds (consonants produced using the lips) (Yang et al., 2019). This area of the brain is associated with sound-symbolic processing in adults. However, this effect was only observed in babies for trials containing the auditory stimulus /moma/ paired with a smooth, rounded shape. Mismatching trials and matching trials containing the auditory stimulus /kiki/ paired with a jagged, spiky shape did not elicit significantly right lateralised activity. Brain areas supporting multimodal integration of sound-symbolic information, vital to acquiring spoken language, are still developing at 11 months (Yang et al., 2019). The perceptual salience of /m/ along with its motoric accessibility and popularity as a Vocal Motor Scheme may result in sound symbolism associated with /m/ being acquired earlier than for /k/ (McCune & Vihman, 1987; DePaolis et al., 2013; Majorano et al., 2014). Indeed a meta-analysis of 11 sound symbolism studies involving 4-38-month-olds found evidence that sound-symbolism recognition for /k/ emerges later in childhood that sound-symbolism recognition for another labial sound /b/ (Fort et al., 2018)

The newborn brain is also dissimilar to the adult brain at the cellular level, containing two to three times as many synapses as the adult brain (Paterson et al., 2006). During infancy, neurons and synapses are selectively preserved (through repeated activation and myelination), adapted, recruited to one hemisphere or the other, or pruned depending on their utility and suitability for performing functions that the individual needs (Johnson, 2000; Paterson et al., 2006; Mareschal et al., 2007; Petanjek et al., 2011). Task-dependent neuronal networks involved in a given behaviour may be much more widely dispersed, degenerate, and redundant in young babies before participating functions become localised to the hemisphere and area where they typically reside in adults (i.e., before they lateralise, regionalise, and localise) (Paterson et al., 2006; Mareschal et al., 2007). That is, neurons that share similar information processing biases may be situated more distally in the developing brain than in the adult brain, a single group of neurons may perform multiple functions, and multiple overlapping networks of neurons may support separate instances of what might appear to the adult observer to be a single behaviour. Before neural networks specialise, one network may perform many functions (Mareschal et al., 2007). Through repeated cyclical interactions between the baby, the environment, and the self, functions are recruited to the specific sets of neurons that handle them most efficiently and some of these neurons may migrate towards one another to optimise transmission of information through networks in which they participate (Thelen & Smith, 1994; Mareschal et al., 2007; Tierney & Nelson, 2009; Paredes et al., 2016).

Whilst the majority of neuronal migration takes place before birth, some neurogenesis and migration continues post-natally, declining rapidly over the first year and ceasing around age 2[29] (Paredes et al., 2016). Having recruited a function, neuronal networks become progressively strengthened through myelination with each repeated activation (Van Der Knaap & Valk, 1990; Su et al., 2008). Sensorimotor processing areas of the brain contain limited myelin at the time of birth and myelination continues relatively rapidly in these areas after birth, influenced by the baby's individual sensory experiences (Van Der Knaap & Valk, 1990; Su et al., 2008). At birth, myelin is not present in other areas involved in language processing including the inferior frontal, superior temporal, or angular gyri, developing later and more slowly in these areas (Su et al., 2008). Nonetheless, myelination of all parts of the brain is most rapid during the first 18-36 months of life, with most major changes in the frontal, parietal, and occipital cortices occurring between 8-12 months, though some myelination may continue at a slower rate for as much as 30 years (Van Der Knaap & Valk, 1990; Hayakawa et al., 1991; Paterson et al., 2006; Su et al., 2008; Williamson & Lyons, 2018; Grotheer et al., 2022).

The networks supporting phonological development and language acquisition are multiple and highly complex, and the formation of networks supporting higher cognitive functions associated with language is dependent on the earlier formation of networks supporting less complex, more generalised types of movement and cognition (Thelen & Smith, 1994; Paterson et al., 2006; Mareschal et al., 2007; Tierney & Nelson, 2009). For example, word acquisition is dependent on progressive integration and interactive specialisation between the capacities for sensory perception, pattern-spotting (e.g., categorical perception, sound-meaning pairings), sensorimotor mapping, neuromuscular control, rhythmic development, understanding of cause and effect, joint attention, symbolic reference, and communicative intent amongst others (Thelen & Smith, 1994; Ejiri, 1998; Johnson, 2000; Mundy et al., 2000; Wolpert, 2003; Keren-Portnoy et al., 2005; Mareschal et al., 2007; Tierney & Nelson, 2009; Ghazanfar & Takahashi, 2014; Vihman, 2014). These capacities emerge asynchronously and are themselves heterogenous with regard to their original functions and relative complexity. We have seen already how some of these 'less complex' capacities themselves emerge out of even earlier capacities.

---

[29] There is some evidence of neurogenesis and migration in adults following brain injury, though this may be driven by different mechanisms to that in babies and toddlers (Kaneko et al., 2017).

Joint attention represents another case in point showing post-natal cerebral specialisation (Thelen & Smith, 1994; Tierney & Nelson, 2009). Joint attention is a vital generalised socio-cognitive capacity for language acquisition but, like syntax, semantics, or pragmatics, it is not necessarily involved in babble (Vihman et al., 1985; Paterson et al., 2006). From 14 months, toddlers show increasingly left lateralised neural activity when responding to extrinsic invitations to engage in joint attention (Mundy et al., 2000). When initiating joint attention themselves, 18-month-olds, but not 14-month-olds, show increased activation in the left frontal cortex and bilateral medial brain areas (Mundy et al., 2000). Initiating joint attention relies on the pre-existing capacity for engaging in joint attention. This, in turn, relies on the ability to direct one's own attention intentionally, typically emerging between 3-6 months, though individual differences in developmental rate and trajectory may affect when joint attentional ability emerges in relation to chronological age and other developmental milestones (Johnson et al., 1991, cited by Mundy et al., 2000, p. 334).

Children, adolescents, and adults

According to fMRI research, lateralisation and localisation of language-related functions may continue gradually and asynchronously for several years after children begin to acquire language and possibly for much longer (Weiss et al., 2008; Rosselli et al., 2014; Olulade et al., 2020; Wang et al., 2021). Five-to-six-year-olds show left lateralised activity in the superior temporal gyrus (BA 41 and 42) and the left supramarginal gyrus[30] (BA 40) during tasks involving phonological processing, whilst semantic processing is associated with left lateralised activity in the left medial temporal gyrus[31] (BA 21) (Weiss et al., 2018). Adults also show left lateralised activity in the left anterior frontal gyrus in speech processing, but children do not show comparable frontal activation until 7-8 years of age (Weiss et al., 2018; Wang et al., 2021). From 7-8 years, phonological processing elicits left lateralised activity in the posterior dorsal inferior frontal gyrus, whilst semantic processing generates left lateralised activity in the anterior ventral and posterior medial temporal gyrus (Wang et al., 2021). Four-to-six-year-olds also show significant activation in the right hemisphere homologues of certain language processing areas – in particular, the right inferior frontal

---

[30] (involved in phonological processing in reading and writing and emotional processing in adults)

[31] (involved in semantic processing during reading and facial and emotional processing

gyrus and the right superior temporal cortex – when making semantic decisions (Olulade et al., 2020). Activity in these right-hemispheric networks declines non-linearly between 6-29 years of age though, as noted earlier in this chapter, some right-hemispheric structures like the right posterior superior temporal gyrus retain a role in language processing throughout the lifespan (Olulade et al., 2020, Harpaz et al.,2009; Binder et al., 2017).

Olulade et al. (2020, p. 24381) suggest that right hemisphere involvement in early speech and language processing may indicate that children first acquire very "basic" representations of speech sounds and language that are initially represented dispersedly or bilaterally and that these differentiate or migrate to the left or right hemisphere with time and repeated experience. Rather than being left lateralised from birth, evidence reviewed by Rosselli et al. (2014) suggests that language-related functions may be represented more bilaterally in children than in adults and that productive and perceptual experience throughout infancy and childhood may be instrumental in the leftward shift of language-related functions. This aligns with the finding that task-dependent networks involved in other, non-linguistic complex functions are also organised more bilaterally and dispersedly in babies than in older children and adults (Paterson et al., 2006). Some synaptic changes, including remodelling and pruning, may continue into the third decade of life, resulting in increasingly more gradual and more focal cerebral specialisation (Petanjek et al., 2011). Far from being purely genetically predetermined and established before birth, lateralisation for language-related functions may continue into childhood and further within-hemisphere cerebral specialisation affecting language processing may continue some time into adulthood.

Differences between the findings of these authors (Paterson et al., 2006; Petanjek et al., 2011; Rosselli et al, 2014; Weiss et al., 2018; Olulade et al., 2020; Wang et al., 2021) and those of Paquette et al. (2015) and Hodgson et al. (2016) may be explained by methodological differences between their studies. Some studies reviewed here examined perception, whilst Paquette et al. (2015) and Hodgson et al. (2016) examined production. However, since perceptual capacities emerge earlier than productive capacities (Vihman, 2014), this difference would not explain why production appears to be stably left lateralised by age 3 but perception does not. Alternatively, the discrepancy may result from differences in the neuroimaging techniques used.

fTCD measures the velocity at which blood flows through the major arteries of the brain – in Hodgson et al.'s study, the mesial cerebral artery. This works on the principle that increased neural activity requires increased blood supply, so blood must flow to activated brain areas at a higher velocity (Deppe et al., 2000). fTCD is therefore limited to providing global information about which hemisphere of the brain requires most blood and over what timeframe (Purkayastha & Sorond, 2013). fMRI examines dynamic changes in the oxygen content of haemoglobin in the blood supplied to different areas of the brain (Deppe et al., 2000). Activated brain areas require more blood because they require the oxygen that it carries.

fMRI maps the specific areas in which activation occurs to within an area of 500µm to 3-4mm depending on the specifications of the apparatus used (Glover, 2011). Weiss et al. and Wang et al. may have captured subtler or more localised changes in lateralisation not detectable by fTCD. Whilst NIRS (as used by Paquette et al.), like fMRI, images specific areas of neural activation, NIRS is more vulnerable to noise, has lower spatial resolution, and is most sensitive to the areas of the brain closest to scalp meaning that it may not provide images of the same depth, accuracy, and granularity as fMRI (Cui et al., 2011; Doi et al., 2013). Thirdly, and critically, Paquette et al. and Hodgson et al.'s studies did not involve children under 3, meaning that the extent to which their findings can be applied to younger toddlers, babies, or foetuses is limited. The first years of life are a time of very rapid, asynchronous physiological and phonological development and change (Van Der Knaap & Valk, 1990; Bogolepova & Malofeeva, 2001; Cheour et al., 2002; Dehaene-Lambertz et al., 2002; Mareschal et al., 2007; Su et al., 2008; Perani et al., 2011; Yang et al., 2019). Whilst early left lateralisation associated with phonetic, phonological, or linguistic processing may seem to indicate genetic determination, a wealth of other research, some of which is discussed here, suggests that experience also plays a critical role.

Summary: left lateralisation as an indicator of linguistic-ness

Left lateralisation, whether genetically or experientially determined or both, is not necessarily an indicator of linguistic-ness. Asymmetries in homologues of classical language areas in humans are also found in other non-human primates who do not possess language, and these asymmetries are correlated neither tightly nor exclusively with language proficiency or lateralisation in young or adult humans (Corballis, 2013). Other capacities like co-ordinating and interpreting complex oral and manual movements; retrieving information from memory; emotional, facial, and general auditory processing; and social, logical, mathematical, and aspects of musical cognition are also left lateralised or involve structures in the left hemisphere (Thelen & Smith, 1994; Binder et al., 1996; Jessen et al., 1999; Griffiths & Warren, 2002; Leech et al., 2012; Leech & Sharp, 2014; Deen et al., 2015; Newman et al., 2015; Amalric & Dehaene, 2016; Elmer et al., 2016; Binder, 2017; Hodgson et al., 2021). Additionally, left-hemispheric structures involved in language processing in adults support other, earlier emerging functions in babies and children. For instance, the left superior temporal gyrus and left inferior frontal gyrus show activation during face processing in babies (Paterson et al., 2006). These findings, once again, support the idea that language processing involves task-specific networks of heterogenous, domain-general parts of the brain rather than some language-specific part of the brain (Petitto & Marentette, 1991; Dehaene-Lambertz et al., 2002; Holowka & Petitto, 2002a; Petitto et al., 2004; Dehaene-Lambertz et al., 2010).

The idea that left lateralisation indicates inherent linguistic-ness may be based on a logical fallacy: it has been argued that because language is left lateralised and language is unarguably linguistic, the fact that babble shares this property tells us that babble is also linguistic (Holowka & Petitto, 2002a). This line of reasoning may result from the theoretical approach adopted by Petitto and colleagues, who have characterised babble as babies' immature attempts to copy language and as evidence of the existence of an innate "brain-based language capacity" (Petitto & Marentette, 1991, p.1493; see also Holowka & Petitto, 2002a and Petitto et al., 2004).

Thelen and Smith (1994) have warned against teleological theories of development. Positioning adult cognition or behaviour as a pre-determined endpoint and characterising development as movement between a series of stepping-stone-like stages allows the researcher to *describe* a developmental journey but not to *explain* it. Teleological theories of development are limited in their ability to account for *how* skills and behaviours emerge and develop asynchronously and non-linearly and, in some cases, must resort to tautological reasoning e.g., that language is left lateralised and so left lateralisation means that something is linguistic (see Petitto and colleagues, 1991; 2002a, 2004). The Old Parts, New

Machine hypothesis, drawing on the principles of evolution and Dynamic Systems Theory (Thelen & Smith, 1994), offers an alternative conceptualisation of phonological development and underlying neural changes in development as a collection of asynchronous and heterogenous phases, movement between which is motivated by interacting biological, cognitive, and environmental factors and constrained by a child's past actions and experiences. Evidence from infant and child laterality research discussed in this section supports this more dynamic view of development and runs counter to claims made by Witelson and Pallie (1973), Paquette et al. (2015) and Petitto and colleagues (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004).

Taking a view of babble and of language as complex heterogenous dynamic systems, it is not clear that babble is linguistic so much as it is articulatory-phonetic or, once systematicity emerges, phonological. Whilst sounding more speech-like than the behaviours that precede it, the resemblance to language may be superficial, since babble does not require any social, syntactic, pragmatic, referential, symbolic, or paralinguistic processing (Oller, 1980). Working, as the present thesis does, from a non-teleological, developmentally-oriented perspective, it is suggested here instead that language may come to be left lateralised if the generalised phonological and cognitive functions that support it are left lateralised some time before babies acquire their first words. It is possible that cases of overall atypical lateralisation for language may arise when some percentage of the functions supporting language come to be atypically lateralised either during pre-natal development or before the time when the first words emerge. Viewed in this way, Holowka and Petitto's (2002a) findings may tell us more about the ontogeny and phylogeny of language than about the underlying nature of babble.

The laterality of vocal communicative functions and handedness

Most people are more proficient with one hand than the other. Whilst both hands may be almost equally good at receptive functions like catching or deflecting objects, the dominant hand is generally better at instantiating planned, original gross and fine motor movements like self-feeding or sewing (Corballis, 2017). In bimanual tasks, the non-dominant hand usually provides balance and stability e.g., in cricket bowling or opening a jar (Serrien et al., 2006). Handedness is reflected in cortical asymmetries. For example, the brains of right-handers show discrete left-hemispheric motor cortical areas for controlling each digit on the right hand and the right wrist, whereas control of the digits on the left hand and the left wrist are represented overlappingly in the right motor cortex (Serrien et al., 2006).

Historically, handedness has been considered a uniquely human trait (Vallortigara et al., 1999; Corballis, 2014b). This, along with the findings that bodily movement is controlled contralaterally; that the majority of people are right-handed; and that language is left lateralised, has given rise to the idea that language lateralisation and handedness may share a common determinant (Vallortigara et al., 1999; Corballis, 2014b). This section will briefly outline why the present thesis does not examine handedness alongside the laterality of babble. The primary reason is that, while the first signs of hand preference may appear between 6-13 months, this is not the case for all babies, and the direction of handedness may not be established until age 2-3 years, with continuing changes in strength spanning childhood (McManus et al., 1988; Ferre et al., 2010). This falls outside the proposed term of commitment for the families involved in this study. The second reason is that the relationship between handedness and the laterality of language and vocal behaviour is neither wholly inextricable nor wholly straightforward (see Knecht et al., 2000; Häberling & Corballis, 2016; Häberling et al., 2016).

Handedness and language

Some research evidence supports the idea that handedness and the laterality of language are related. For example, some behavioural and neuroimaging (fTCD) paradigms have found similar patterns of laterality for language and manual manipulation and communicative pointing in adults, toddlers, and children (Esseily et al., 2011; Cochet & Vauclair, 2012; Hodgson et al., 2016; Hodgson et al., 2021). Both language and handedness are predominantly left lateralised by age 3, and individual children with atypical language laterality show more inter-individual variation in.the proficiency gap between their dominant and non-dominant hand compared with typically developing children (Hodgson et al., 2016). That is, . Fourteen-month-olds who show a left hand preference for pointing may have smaller receptive and productive than right-handed pointers, and those who do not point at all may have smaller vocabularies still (Esseily et al., 2011).

Two major theories have strengthened the perceived relationship between language laterality and handedness. Annett's (1996) *Right Shift theory* posits that a single gene is responsible for both. In individuals with the dominant RS+ allele, this gene is suggested to force language and handedness to shift leftward by suppressing activity in the right hemisphere homologues of language and hand control areas. Individuals with the recessive RS- allele, however, have a 1:1 chance of atypical lateralisation and/or left-handedness. Whether a single gene could effect this change is debated, and other authors envisage a greater role for individual experience (Provins, 1997; Sommer & Kahn, 2003). Alternatively, *Gestural theory* suggests that spoken language is left lateralised because it evolved out of right-handed gesturing, which in turn evolved out of left lateralised non-linguistic communicative behaviours in our ancestors (e.g., Corballis, 2003). Gestural theories have attracted criticism centring around plausibility and parsimony (Cook, 2003; MacNeilage, 2003) and Corballis' more recent publications reflect a change in stance (Corballis, 2008; Häberling & Corballis, 2016; Häberling et al., 2016).

However, handedness and language are only weakly correlated (e.g., see Knecht et al., 2000). A robust, tendency towards left laterality for language-related functions is seen across left- and right-handers, with only a minority showing atypically lateralised language-related functions (see Table 3 below). Handedness research has established that people point with their dominant hand the majority of the time (Cochet & Vauclair, 2012). A higher proportion of adults exhibit a right hand preference in *non*-communicative pointing (87.4%) than in communicative pointing (77.2-81.1%) (Cochet & Vauclair, 2012). Whilst 68% of 14-month-old toddlers communicative points may be right-handed, 63% toddlers also exhibit a right hand preference for grasping objects (Esseily et al., 2011).

fMRI research has identified three distinct but partially overlapping left lateralised neuronal networks that support language, gesture, and handedness, which may have arisen through progressive differentiation of the multimodal, domain-general primate mirror system during hominin evolution (Häberling & Corballis, 2016; Häberling et al., 2016). While the left medial temporal gyrus and the pars opercularis and pars triangularis (BA 44 and 45) in the left inferior frontal gyrus participate in the networks for both handedness and language, the network underlying handedness occupies more of the parietal cortex while the language network shows more frontal and temporal cortical involvement (Häberling & Corballis, 2016; Häberling et al., 2016). Whilst activity in the hand motor cortical area is seen during language processing, this is relatively weak, and may be more a case of spreading activation through overlapping neuronal networks rather than focal co-activation (Salmelin & Sams, 2002; Saarinen et al., 2005; Morillon et al., 2010). Spreading activation may offer some insight into why some babies begin to show hand preferences in unimanual and bimanual hand movements in the absence of language close to the times at which singleton, reduplicated, and variegated babble emerge (Ramsay, 1980, 1984). Prelinguistic specialisation in the left hemisphere resulting from repeated practice of planning and articulating sequential vocalisations may have the side effect of also progressively strengthening the partially overlapping left-hemispheric network responsible for hand movement.

% people left lateralised for language	Graves et al. (1982)	Knecht et al. (2000)	Nenert et al. (2017)	Van Der Haegen and Brysbaert (2018)
Right-handers	72.6-77.4%	90-96%	93-95%	--
Ambidexters	--	85%	75-78%	--
Left-handers	68%	73-78%	75-78%	80%

Table 3 Percentage of right-handers, left-handers, and ambidexters showing left lateralisation for language in four studies of laterality and handedness.

Handedness and the laterality of communicative vocalisation in non-human animals

The idea that handedness is inextricably linked to language are further weakened by observations that unilateral limb preferences are seen across the order *Bilateria*, in species as diverse as primates, rodents, birds, and amphibians (Vallortigara et al., 1999). Right-handedness is seen in approximately 65% of chimpanzees, with higher proportions in wild than captive chimpanzees, suggesting that right-handedness is typical of both human and non-human great apes (Hopkins & Cantalupo, 2003; Hopkins & Cantero, 2003). In particular, chimpanzees exhibit a right hand preference for gesturing, which is exaggerated when paired with a communicative vocalisation (Hopkins & Cantalupo, 2003). Grey mouse lemurs exhibit hand dominance, which may be associated with lateralisation for processing species-typical calls, but only weakly if at all (Scheumann & Zimmermann, 2008).

These findings may seem superficially to offer support for Gestural theory. However, most haplorrhine primates are facially mobile and both strepsirrhine and haplorrhine primates are highly voluble and, and possess communicative systems of varying complexity, made up of differentiated vocalisation types with contexts-specific meanings, which are transmissible over great distances (Masataka & Fujita, 1989; Arcadi, 2003; Holloway, 2003; Arnold & Zuberbühler, 2006; Ouattara et al., 2009; Schel et al., 2013b). The availability of vocal communicative systems like these across primate species including humans suggests that unimanual gesturing was not necessarily an essential part of language evolution but rather may only serve an optional, amplificatory role and only when individuals are in close proximity and can see one another (Bradshaw, 2003). Chimpanzees exhibit the ability to learn, represent, and modify vocal signals to instigate and prolong joint attention and social bonding episodes with both conspecifics and heterospecifics (Hostetter et al., 2001; Slocombe & Zuberbühler, 2007; Losin et al., 2008; Slocombe et al., 2010b; Wallez et al., 2012; Russell et al., 2013; Schel et al., 2013a; Fedurek et al., 2015a). It seems likely, then, that the capacity for understanding, flexibly reproducing, and innovating vocal communicative acts likely evolved separately from handedness in our early primate ancestors.

Alternative explanations for the evolution of handedness

Handedness may be more strongly associated with bipedalism than with language evolution. Gorillas and orangutans, for example, exhibit a right-hand preference when standing on two legs but no such preference when handling objects or gesturing whilst standing or movement quadrupedally (Corbetta, 2003; Pedersen & Vereijken, 2003). Early bipedal human ancestors exhibit signs of right-handedness for throwing and tool use – skills that involve complex cognition but that long pre-date language (Wolpert, 2003). The position and mobility of *Homo erectus’* (1.5-1.6 MYA) arm and shoulder joint is compatible with tool use and analysis of the imprints left by their musculature suggests that they were proficient at throwing (Roach & Richmond, 2015). *Nariokotome boy*, a *Homo ergaster* skeleton (dated to 1.5-1.9 MYA), exhibits a longer right than left ulna and patterning on the clavicles indicative of a more well-developed right than left deltoid muscle, reflecting the anatomy of right-handed modern humans (Schulter-Ellis, 1980; Walker & Leakey, 1993; Beaton, 2003). The teeth of one *Homo habilis* skull (1.8 MYA) show a pattern of repeated accidental damage consistent with right-handed but not -left-handed tool us (Frayer et al., 2016). However, the first stone tools date back to 2-2.6 MYA (Beaton, 2003) and bipedalism to *Sahelanthropus tchadensis* (4-7 MYA) (Brunet et al., 2002; Holloway, 2003) meaning that hominin handedness may pre-date language by several million years.

The tendency towards right-handedness may be explained not by a deterministic relationship with language but by the typical propensity of the left hemisphere for representing complex sequential actions (Häberling & Corballis, 2016; Häberling et al., 2016). Only manual tasks that, like language, involve sequential movement show a consistent correlation with the lateralisation of language (Hodgson et al., 2021). Direct contralateral projections between the left hemisphere and the right hand may create a species-level bias towards right-handedness for sequential tasks like writing or performing multi-step manual actions. In the absence of direct ipsilateral projections, innervating the left hand for activities like these would involve indirect transmission of activation from the left motor cortex to the left hand via the corpus callosum (Hodgson et al., 2016). Gestures and non-communicative, non-sequential manual movements may be processible in more global, holistic ways, achievable by more right lateralised or bilateral neuronal networks (Woll & Sieratzki, 2003; Häberling & Corballis, 2016; Häberling et al., 2016).

If handedness is task- and physiology-dependent, this may suggest a greater role for genetics than for individual experience (Annett, 1996). Recent research does not indicate a single gene theory like Right Shift theory (Annett, 1996) but rather, a case of *partial pleiotropy* (Ocklenburg et al., 2014). That is, language lateralisation and handedness may be determined by overlapping groups of genes, which may affect each trait in different ways. FOXP2, associated with speech production, has recently been implicated in handedness (Crespi et al., 2017) and LRRMT1 and PCSK6, classically associated with handedness, have been implicated in neurolinguistic differences in schizophrenia and dyslexia (Corballis, 2014b). Differences in the expressions of COMT, also typically associated with handedness, have been linked to the rate and trajectory of development of the posterior language and frontal attentional areas of the brain in 6-10-year-olds (Sugiura et al., 2017). As far as the author of this thesis is aware, no association with language has been found for other genes implicated in handedness, including APOE, AR, LMO4, MT2A, or STK35 (Ocklenburg et al., 2014; Schmitz et al., 2018). This suggests that there is only partial rather than complete overlap between the genetic variants that determine lateralisationfor language and lateralisation for handedness and this may explain why these two traits are only weakly correlated (e.g., see Knecht et al., 2000; Nenert et al., 2017).

Summary

So then, whilst represented proximally in the brain, vocal and communicative behaviours and handedness are only weakly associated. Left lateralisation for language and handedness likely evolved separately and are supported by overlapping but critically distinct neuronal networks. Future research could investigate whether babies showing signs of possible future right- and left-handedness show similar trends in the laterality of vocal behaviours as right- and left-handed adults do for language (Graves et al., 1982; Wyler et al., 1987; Knecht et al., 2000). However, since the present study builds on Holowka and Petitto's (2002) study, which did not examine handedness, these questions fall outside its scope.

Concluding remarks

The evidence and the hypotheses

The evidence discussed in this chapter aligns most closely with a conceptualisation of language in adults, communicative behaviour in non-humans, and pre-linguistic phonological behaviour in babies, as distinct complex dynamic systems composed of multiple heterogenous, asynchronously changing subsystems (i.e., behaviours, skills, capacities) which may become increasingly sophisticated over developmental and evolutionary time given, sufficient intrinsic or extrinsic impetus. This is the position taken by the Old Parts, New Machine hypothesis.

Early conceptualisations of the Motoric hypothesis that posited discontinuity between babble and language are further weakened by the finding that, as in behavioural development, parts of the brain supporting phonological development and language acquisition show continuous, gradual, and non-linear specialisation during infancy and childhood and, in some cases, into adulthood. Evidence that ingestion and vocalisation are supported by distinct neuronal networks aligns with behavioural and acoustic evidence in posing a problem for later conceptualisations of the Motoric hypothesis positing that the rhythms of language emerge out of ingestive cyclicities, including Frame/Content theory (Davis & MacNeilage, 1995).

Nor is the Linguistic hypothesis well-supported by the evidence discussed here. Many domain-general structures in the brain are also left lateralised, (e.g., those supporting complex sequential mouth and hand movement, mathematical reasoning, general cognition, attention, memory, social cognition, and even certain aspects of musical cognition). This calls into question whether left lateralisation constitutes reliable evidence to support the claim that babble is linguistic. Evidence discussed in this chapter suggests that sequential vocalisation, both meaningless and meaningful, in humans and non-humans, may be left lateralised quite independently of language. Additionally, whilst some degree of laterality may be constrained by genetically encoded information-processing biases, laterality is not complete or stable by 12 months of age, much less at the time when babies first begin to babble. There seems to be good reason then, to expect that babble might becoming more strongly or more consistently left lateralised with time and practice.

The Old Parts, New Machine hypothesis argues that babble is not linguistic, but rather an endogenously emerging rhythmic-motoric vocal behaviour. Under this hypothesis, babble comes to take on some socio-cognitive significance only once babies accumulate enough articulatory and perceptual experience to recognise articulatory and acoustic similarities between other people’s vocalisations and their own. This accumulated skill and experience then becomes coupled to the impetus to behave like those other people around them. Earlier similarly non-linguistic manifestations of this impetus may be seen in babies’ and caregivers’ reciprocal imitation of clapping, vocalic ‘singing’, and ‘blowing raspberries’ (see also chimpanzees!). Both babble and language are supported by domain-general parts of the brain that are involved in producing and perceiving vocal sounds but that are also involved in other non-linguistic behaviours. Unlike babble, language also involves a myriad of other capacities and brain areas. It is argued here that what makes language linguistic is not purely its articulatory and phonetic properties, but all of the specific interactions between the networks of brain areas that support functions involved in language processing. Emerging in the absence of these other capacities, it is argued here that babble should not be considered linguistic.

The methodology

This chapter has also considered the power and suitability of orofacial asymmetry analysis as a tool for analysing laterality. Graves and Landis (1990) comment that oral asymmetries are not only manually measurable but are also sometimes visible to the naked eye in day-to-day conversation. Orofacial asymmetry analysis makes these visible differences quantifiable. Importantly, orofacial asymmetry analysis is also sensitive to subtler inter- and intra-individual variations in the degree and direction of asymmetry of the type that have since been identified in neuroimaging and neuroanatomical research and that are less perceptible to the naked eye. Non-invasive methods for analysing laterality have significant strengths. They are cost-effective and can be used to analyse laterality in naturalistic behaviour across the lifespan and across species. This kind of work can give us valuable insights into the ontogeny and phylogeny of language.

However, because most human studies to date have used very different methods for measuring orofacial asymmetry, it is difficult to determine to what extent the differences in the findings of existing studies can be directly compared. In the late 1990's, Hook-Costigan and Rogers (1998) developed a new and more rigorous, objective, and sensitive method for measuring orofacial asymmetry (see also Fernández-Carriba et al., 2002a; Fernández-Carriba et al., 2002b; Schuetze & Reid, 2005; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). This method can be used with participants across ages and primate species and can be applied to naturalistic as well as experimental data. The present study takes this method and develops it by addition of a further exclusion criterion and a further measure for accurately identifying the midline of the mouth in human babies. For the purposes of comparability, it may be informative for future research to ascertain whether the findings of the adult orofacial asymmetry research outlined here (Graves et al., 1982; Wolf & Goodale, 1987; Wyler et al., 1987; Hausmann et al., 1998) are replicated with this more rigorous method.

Very little research to date has looked at orofacial asymmetry in babies. Holowka and Petitto's (2002a) findings raise some interesting questions. This study was cross-sectional and drew data from babies aged from 5-11 months, matched for the time when reduplicated babble emerged.

Firstly, it would informative to know whether left lateralisation in babble is fairly consistently strong and stable over multiple instances of production or whether it is initially more variable, becoming more consistent with time or articulatory practice. Changes like these might indicate increases cerebral specialisation during babble of the type discussed by Ramsay (1980, 1984) in relation to the emergence of reduplicated and variegated babble. Secondly, it is not yet known whether the left lateralisation that the Holowka and Petitto (2002a) observed remains stable over developmental time, whether it can be observed from the time when babble emerges, or whether it emerges during singleton babble, or increases or declines with age or articulatory practice.

If babble is stably left lateralised from emergence, this may indicate a dominant role for genetic pre-determination *in utero* or experience-dependent cerebral specialisation resulting from previous productive or perceptual experiences. If babble shows right lateralisation or no clear lateralisation at the time of emergence, with left lateralisation emerging across the period of data collection, this may indicate a predominant role for experience-dependent processes that are specific to canonical syllable production. If babble is somewhat left lateralised or shows great variability in lateralisation across utterances close to the time of emergence, becoming more strongly and/or consistently left lateralised across the data collection period, this may indicate a combined, interactional role for genetics and experience. That is, the left hemisphere may recruit babble because it is the more strongly predisposed to producing sequential vocalisations than the right, and this favouring may, in turn, strengthen the left hemisphere's capacity for babble. The evidence discussed in Chapters 1-3 of this thesis suggests that this third possibility is most likely. During babble, babies begin to exhibit more systematic, though still highly idiosyncratic, vocal behaviour. It seems likely that the emergence and development of systematic vocal behaviours should be associated with more systematically organised neural activity (Ramsay, 1980, 1984; Vihman, 2014; Sussman, 2015). Research reviewed in this chapter suggests that the left hemisphere houses a greater proportion of domain-general neurons that are biased towards processing complex, sequential information, whilst the right hemisphere recruits neurons biased towards global, holistic, gestalt-like processing (Sussman, 2015). The laterality of babble may be more a product of these domain-general information processing biases than of babble being inherently linguistic or language-like in some way.

Thirdly, it would be interesting to know whether any changes in laterality that might be observed over time are associated with any advances in phonological skill. Given the key theoretical assumption of Dynamic Systems Theory, that different skills and capacities develop asynchronously, it seems likely that some of the babies involved in this study might each have attained quite different collections of social, emotional, and motor milestones by the times of testing, and perhaps also some different vocal milestones. For example, since Holowka and Petitto (2002a, 200b) began testing only once *reduplicated* babble had emerged, it is quite possible that some of the older babies in the sample – those aged 10, 11, and 12 months – might already have settled on one or two vocal motor schemes (McCune & Vihman, 1987) in their singleton babble, whilst younger babies may not yet have developed any such production preference. Finding change in the laterality of babble associated with articulatory experience might indicate that articulatory experience brings about neural reorganisation. This would have implications for the Linguistic hypothesis, which posits that left lateralisation is biologically rather than experientlially determined

The present study aims to address these gaps in the research and, in doing so, hopes to contribute to our understanding of how phonology, sequential movement, and language come to be left lateralised. The following chapter will now outline the method used in the present study to seek answers to these questions.

Chapter 4 Data and methods

This chapter will recapitulate the three research questions emerging from the literature that are explored in the present study. It will then outline the data collected and the methodological decisions undertaken to address these questions. Measures taken to ensure reliability are also outlined.

Introduction

Holowka and Petitto (2002a) found evidence of differential hemispheric specialisation for babble, smiles, and non-babble vocalisations in 5-12-month-olds. They found babble to be left lateralised and smiles to be right lateralised. Similarly, adult laterality research has found speech and non-speech mouth movements to be left lateralised whilst emotional facial expressions show right laterality (Graves et al., 1982; Wolf & Goodale, 1987; Wyler et al., 1987; Wylie & Goodale, 1988). Laterality research with primates has also found evidence of analogous hemispheric specialisation (e.g., Losin et al., 2008). Sequential or artificial, emotionally neutral vocalisations have been found to show left hemisphere dominance, whilst species-typical emotional vocalisations show right hemisphere dominance (Hook-Costigan & Rogers, 1998; Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). Whilst Holowka and Petitto (2002a) and Graves and colleagues (Graves et al., 1982; Wyler et al., 1897) measured laterality via the angles created by the opening of the lips, these primate studies developed a more sensitive and rigorous method for analysing laterality via orofacial asymmetry. This method – hemimouth area measurement – involves measuring the relative area of the lip opening on each side of the mouth. An adapted version of this method is used in this thesis.

Some asymmetries in the morphology of the brain may be seen from or before the time of birth e.g., the left planum temporale is larger than the right at birth and right inferior frontal gyrus is more differentiated than the left before 28 days of age (Witelson & Pallie, 1973; Chi et al., 1977; Wada, 1977; Bogolepova & Malofeeva, 2001; Reissland et al., 2014). However, the brain continues to develop structurally and functionally after birth and the first year of life is a time of great reorganisation and change in the parts of the brain that later come to support functions related to language and social communication (Van Der Knaap & Valk, 1990; Bogolepova & Malofeeva, 2001; Cheour et al., 2002; Dehaene-Lambertz et al., 2002; Mareschal et al., 2007; Su et al., 2008; Perani et al., 2011; Yang et al., 2019). Some neural reorganisation of these areas and the networks in which they are involved continues long into childhood and adolescence (Weiss et al., 2018; Olulade et al., 2020; Wang et al., 2021). One study with babies and toddlers found evidence that negatively valenced emotional facial expressions continued to become more strongly right lateralised into the second year of life (Schuetze & Reid, 2005). Ramsay (1980, 1984) observed changes in babies' hand movements coinciding with the emergence of more complex subtypes of babble. Ramsay observed the emergence of a right- or left-hand preference in one-handed hand movements coinciding with the emergence reduplicated babble emergence (1984), and a right- or left-hand preferences in in two-handed hand movements coinciding with variegated babble emergence (1980). Ramsay interpreted this as evidence that some underlying change in hemispheric specialisation was taking place. It is possible that this underlying specialisation is restricted to the hand motor cortical areas and is not associated with wider changes neural organisation affecting other nearby brain regions. Recall that handedness is determined by an overlapping but functionally locally distinct neural network to speech (Häberling & Corballis, 2016; Häberling et al., 2016). However, Oxley et al.'s (2014) longitudinal pilot study involving two babies aged 9-18 months found evidence to suggest that mouth movements involved in babble may become more strongly left lateralised around the time when babies attain a Vocal Motor Scheme (Oxley et al., 2014). This may suggest that some process of hemispheric specialisation is also taking place in the parts of the brain responsible for babble during this time like the lip, tongue, and jaw areas of the motor cortex and the inferior frontal gyrus.

Research questions

Upon considering the findings of previous laterality research, a number of questions emerge. The present study aims to address three of them. Firstly, can Holowka and Petitto's

(2002a) findings be replicated using the more sensitive and rigorous method for analysing laterality via orofacial asymmetry developed by primate researchers? That is, will hemimouth area measurement show babble to be left lateralised and smiles to be right lateralised? Secondly, does the laterality of babble show any change with time and accumulated articulatory experience between the time when it emerges and the time when some babies begin to say their first words? Thirdly, if change in laterality is seen during this time, is this change associated with the attainment of any milestones in pre-linguistic development that indicate an increase in systematicity or articulatory skill? The milestones examined here are the attainment of Vocal Motor Schemes; singleton, reduplicated and variegated babble; and monosyllabic, bisyllabic, and polysyllabic babble. These milestones were selected for their salience and relevance both to previous laterality research and to later language acquisition.

Building on previous research

In building on Holowka and Petitto's (2002a) research there are two key methodological decisions that require consideration. These relate to the authors' sampling of data and to the criteria used.

Firstly, Holowka and Petitto (2002a) pooled data from a small sample of infants of different ages and handle these data as homogenous. This decision may have been undertaken because the participating babies were at the same developmental stage, but this is not explained explicitly. The authors state that data was collected from each baby "between the ages of 5 and 12 months according to the age at which each baby first entered the syllabic babble stage" (Holowka & Petitto, 2002a, p. 1). Given the diversity of the chronological ages of babies sampled, it is very likely that some babies had more language exposure, more articulatory experience, and more experience of social interaction, than other babies in the sample. In their supplementary material (Holowka & Petitto, 2002b, p. 1), the authors state that "the exact age of each baby at testing and the age of onset of syllabic babbling are one and the same thing" but it is not stated how it was ensured that babies were recorded on the same day that they first produced syllabic babble e.g., whether the caregivers were asked to carry out the audio-visual recording, whether the family attended a university laboratory immediately or were immediately visited at home by a researcher. It is also unclear whether the audio-visual recordings were taken under naturalistic conditions, or whether any stimulus was used to prompt babies to produce babble, non-babble vocalisations or smiles. The latter may be implied since (Holowka & Petitto, 2002b, p. 1) used the word “testing”, though, again, this is not made explicit.

The second point for consideration relates to the authors’ criteria for babble. Holowka and Petitto (2002a, p.1515) state that their criteria are that babies’ vocalisations a) must contain a “reduced subset” of the consonants of adult language; b) must feature reduplicated syllables; and c) must be produced without meaning or reference in order to be judged as babble. This second criterion diverges from more widely accepted criteria for babble which include singleton babble and variegated babble as well as reduplicated babble (Oller, 1980; Elbers, 1982; Elbers & Ton, 1985; Oller & Eilers, 1988; Oller, 2000). It is not explained why these types of canonical babble are excluded here. Petitto et al. (2004) and Petitto and Marentette (1991) also use these reduced criteria, excluding singleton and variegated babble, in their studies examining manual babble in hearing and hearing impaired, speech-exposed and sign-exposed babies.

Petitto and colleagues (1991; 2002a; 2004) cite Elbers (1982), Locke (1983), and Oller and Eilers (1988) as the source for these reduced criteria. In discussing her own baby – Thomas – and his babble development, Elbers (1982) outlines four chronologically ordered stages. During the first stage, Thomas produced only *single babbles* in isolation (e.g., [əbvv], Elbers, 1982, p. 55) before beginning to produce *repetitive babble* (e.g., [bəbəbəbəbəbə], p. 55) in the second stage. In the third stage, Thomas began to *concatenate* discrete repetitive babble utterances into longer sequences, which contained consonants with different places and manners of articulation (e.g., [bəgəbəgəbəgəbək] p. 55). In the fourth stage, Thomas began to combine different consonants together within single babble utterances (e.g., [gɑŋgwɑŋgəb ŋəməkə əwbɑbuməkə] p. 55). Stages one, two, and four bear close resemblance to Oller’s (1980) descriptions of singleton, reduplicated, and variegated babble. In the introduction to their paper on the role of auditory perception in syllable production in hearing and hearing-impaired babies, Oller and Eilers (1988, p. 442) briefly outline Oller’s (1980) stage-by-stage explanation of how babble emerges from previous vocal behaviours, and they state that canonical babble is “characterised by production of reduplicated utterances”. Oller and Eilers (1988) do not discuss singleton or variegated babble here, but nor do they say that canonical babble excludes singleton and variegated babble. In his 2000 book, *The Emergence of the Speech Capacity*, Oller reiterates his 1980 criteria for canonical babble stating “[i]n the *Canonical stage* [sic], infants produce closure and opening sequences with normal phonation in well-timed, often repetitive patterns. […] In the most salient patterns of the stage infants produce reduplicated sequences” (p. 65)**.** Here Oller (2000) restates that canonical babble *may* be reduplicated but does not state that this is a necessary requirement. He goes on to provide a spectrograph for a “canonical syllable [tat].” (p. 66). The paper by Oller and Eilers (1988) appears to be the most likely source of Holowka and Petitto’s (2002a) criteria for babble and may account for the mismatch between their criteria and the more widely accepted criteria for canonical babble. This decision may have been undertaken as Oller (2000, p. 65) has pointed out that reduplicated babble constitutes the “most salient” form of canonical babble. This may be easier for caregivers to identify and/or may expedite identification of suitable still-frames for analysis. However, again, it is not stated whether this is the case.

These methodological considerations are raised here since the present study will differ in its handling of these decisions. The present study will begin and end data collection at the same chronological age for each baby and detailed information will be collected on when babies begin to babble, the articulatory form of their babble, and when they develop favoured consonants. As discussed in Chapter 1, the present study will also adopt Oller’s (1980) original criteria for identifying babble. Excluding singleton and variegated babble from babble research may risk mis-representing babble as static or stable in form. Research suggests that babies may progress through these phases in babble as they become more proficient at producing syllables (Oller et al., 1976; Oller, 1980; Ramsay, 1980; Stark, 1980; Elbers, 1982; Ramsay, 1984; Elbers & Ton, 1985; Studdert-Kennedy, 1990; Elbers, 2000; Oller, 2000; Vihman, 2014). It is possible that babies’ growing proficiency at producing babble may be associated with changes in the laterality of babble, as the parts of the brain involved in babble become more specialised during development. It is also possible that moment-to-moment changes in the articulatory gestures involved in one babble utterance compared with the next may be associated with differences in laterality, being supported by slightly different patterns neural activity.

Participants

Babies and their families were recruited through the York BabyLab, which advertises through the parenting-themed Facebook groups York Mumbler and York Nurturing Community. Recruitment was restricted to families living in a ≤6 mile radius of the University, who had babies aged ≤5 months who were born at full term, had no known developmental or auditory problems, and had not yet begun to babble. A sample of 12 babies (5 male, 7 female) were recruited and laterality data from eight of these babies (pseudonyms Freya, Benji, Cameron, Orelia, Fred, Adelaide, Arthur, and Leif) is analysed in the following chapter. Participating families were compensated for their commitment with a £60 participation fee and a York BabyLab t-shirt. Informed consent was obtained from each participating family prior to commencing data collection (see Appendices I-II). Participants were informed that they could withdraw from the study at any time without explanation and would receive a proportion of the full participation fee commensurate with the actual term of their involvement in the study (see Appendix I). None of the families who took part chose to withdraw early. The method for analysing orofacial asymmetry used in this study is particularly time-consuming. The decision was made to limit the number of babies' data to be analysed to the first eight babies from whom data was collected (5 male, 3 female) in order to allow for a more fine-grained longitudinal analysis of the data. These babies were 152-181 days old at the time of the first visit (mean age 162.25 days, SD 10.05 days) and 356-364 days old at the time of the final visit (mean age 359.38 days, SD 2.83 days)

Materials

Audio-visual recordings were created using a Panasonic HDC-TM700 video camera (July-November 2017) and a Canon XA30 video camera (November 2017 onwards). Both video cameras have a high resolution of 1080p, giving high quality images and facilitating fine-grained measurement of the size the opening between the lips. However, the Canon XA30 captures 50 frames per second and generates .mp4 files, whilst the Panasonic HDC-TM700 captures only 25 frames per second and generates files in .mts format, which require time-consuming conversion before analysis is possible. A higher frame rate facilitates more accurate identification of the point at which the mouth opens the widest. These features of the Canon XA30 were considered to be sufficiently advantageous to justify a change in equipment when the Canon XA30 became available for use. A Sennheiser SK 100 Bodypack transmitter and receiver were used for audio capture. The receiver pack was fitted to a LENA waistcoat (see Figure 5), This is a cotton waistcoat with flexible synthetic side panels which fastens at the front with poppers and has a horizontal pocket on the back to house the LENA recording device (https://www.strollerbuzz.com/lena/). These waistcoats are used by the University of York BabyLab and are specially made for use with the LENA recording device and are designed to be comfortable and safe to use with babies. The waistcoat was either worn by the baby or placed as close as possible to the baby in cases where wearing the waistcoat might cause the baby discomfort or distress e.g., during hot weather, or during the time before the baby could sit independently. Testing prior to home visits found a range of one metre or less to give audio data of sufficient quality for coding and transcription.

Figure 5 Image showing a LENA waistcoat. This image is reproduced from https://www.york.ac.uk/language/research/projects/babylab/lena-device/#tab-1

Data

A series of 30-minute audio-visual recordings of naturalistic play between babies and their families were collected. Data collection began at 5 months of age and ended at 12 months to maintain a consistent term of commitment across families and for comparability with Holowka and Petitto's (2002) sampling. Since babble typically emerges at 6-8 months (Oller, 1980; Vihman, 2014), two initial visits were made no more than a fortnight apart starting at 5 months (mean age = 0;5.11, i.e., 162.25 days) with the aim of obtaining a 'baseline' impression of any relevant individual lateral bias in facial structure or mobility. At the end of this second visit, families were asked to contact the researcher to re-commence visits once their baby began to babble. The author explained the criteria for babble used in the present study and provided families with examples of babble compared with non-babble vocalisations. Once babble emerged, families were visited twice per month for the remainder of the data collection period. Babies who began to babble earlier were therefore visited more times than later babblers. An ideal interval of two weeks between visits was maintained where possible. However, in some cases, visits had to be scheduled, postponed, or rescheduled around families' commitments (e.g., GP appointments, baby groups, house moves, family visits or holidays) or unforeseen circumstances (e.g., family illness, or babies feeling too tired, fussy, or hungry to participate). Table 4 details the number of visits made to each baby and the baby's age at each visit.

Visit	Cameron	Benji	Freya	Orelia	Fred	Adelaide	Arthur	Leif
1	159	172	165	154	152	154	181	161
2	171	182	181	160	166	161	186	181
3	224	238	194	216	286	299	333	258
4	227	252	217	226	300	309	348	272
5	237	267	258	251	329	324	362	290
6	258	294	271	282	342	336		301
7	276	308	295	296	356	345		316
8	343	322	307	316		364		331
9	357	337	320	332				358
10		354	336	346				
11		361	357	360				

Table 4 Number of visits made to each baby and the baby's age (in days) at each visit

Analysis and reliability

Transcription and coding for expression

ELAN version 4.9.4 (2016) was used to transcribe and code the audio-visual data and to extract still-frames for analysis. The videos were first coded to identify instances of the three categories of orofacial gesture of interest to the present study: babble, non-babble vocalisations, and smiles or laughs. This coding was used to identify suitable still-frames for analysis. The babble category comprised utterances fitting Oller's (1980) criteria for babble: these utterances contained one or more CV alternation/s produced with adult-like temporal organisation and articulation and without meaning or reference (e.g., [ba], [mamama], [utɪgəʃəw]). The non-babble category contained other vocalisations not fitting these criteria like growls, squeals, sighs, vocalic shouts, burps, sneezes, yawns, coughs, hiccoughs, raspberries, and grunts.

In the initial two baseline visits before babble emergence, only non-babbles and smile/laugh were coded for most babies since they had not yet begun to babble. A small number of marginal babble utterance were identified during transcription for Arthur (n = 5) and Leif (n = 2) during their second baseline visits. These marginal utterances contained some consonants and vowels, but they were not produced with adult-like precision or rhythm. When asked, their caregivers stated that these babies produced vocalisations only rarely.

From the time of babble emergence onwards, all three categories were coded. Babble utterances were also transcribed using the International Phonetic Alphabet (see Appendix IV) to identify the point when babies attained a Vocal Motor Scheme or Schemes (McCune & Vihman, 1987; McCune & Vihman, 2001) (see Appendices V-VI). Vocal Motor Schemes were identified following the procedure outlined by McCune and Vihman (1987; 2001; see also Vihman, 2014). Following transcription, the different consonants produced by each baby in each utterance during each visit were counted. Following McCune & Vihman (1987, 2001) place and manner were treated as distinguishing features but voicing was not (following Macken & Barton, 1980; Stark, 1980; DePaolis et al., 2013). That is, /b/ and /d/ were treated as separate consonants, as were /d/ and /z/, but /t/ and /d/ were grouped together. Consonants that were produced in >10 babble utterances over 3 out of 4 consecutive sessions or consonants that were produced >50 times in a single session were identified as Vocal Motor Schemes. Production of a given consonant with this frequency and consistency is taken to indicate that the baby has an entrenched articulatory routine for producing this consonant and favours this consonant over other less frequently or consistently produced consonants in their repertoire (McCune & Vihman, 1987; McCune & Vihman, 2001; Keren-Portnoy et al., 2010; DePaolis et al., 2011; Vihman et al., 2014). All Vocal Motor Scheme analysis was carried out after the data collected for each child had ended to avoid implicit researcher bias. Following McCune and Vihman (1987, 2001) the first visit in which a Vocal Motor Scheme consonant began to show evidence of being a Vocal Motor Scheme was taken as the *during_vms* visit (i.e., the first of 3-4 visits containing >10 utterances containing the consonant in question, or the first visit containing >50 utterances containing the consonant in question). All sessions before this session were coded as *before_vms* and all visits after this were coded as *post_vms*.

Possible early word forms (imitated and spontaneous) were also transcribed for use in identifying Vocal Motor Schemes. The laterality of these possible word forms was not analysed because there were too few cases to form a category and because the primary focus of the present study was to examine the laterality of babble rather than words. Whilst these possible word forms were similar in phonological form to babble, unlike babble they were produced *with* symbolic reference and/or communicative intent. Given that the neural circuitry involved in auditory-articulatory mapping is distinct from and more strongly left lateralised than the circuity involved in meaningful language processing (Hickok & Poeppel, 2007; Corballis, 2015), possible word forms may be less strongly lateralised than babble. Future research with older babies and toddlers could investigate whether this is, indeed, the case. Table 5 shows the number of visits following babble emergence that fell before and after the emergence of the first Vocal Motor Scheme. Table 6 shows the Vocal Motor Schemes attained by each child and their age at Vocal Motor Scheme attainment.

	Cameron	Benji	Freya	Orelia	Fred	Adelaide	Arthur	Leif
before_vms	3	3	6	0	0	2	1	6
during_vms	1	1	1	1	1	1	1	1
post_vms	3	5	2	8	4	3	2	0

Table 5 Number of visits before_vms, during_vms, and post_vms attainment for each baby.

baby	p/b	t/d	k/g	ʔ	m	n	ŋ	ɸ/β	θ/ð	h	l	j
Freya		0;10.16			0;10.16							
Benji		0;9.21	0;9.21									
Cameron					0;8.15							
Orelia		0;7.2		0;11.9								
Fred		0;9.12										
Adelaide		0;10.20								0;11.13		0;11.13
Arthur		0;10.29						0;10.29	0;10.29		0;11;14	
Leif		0;11.23										
Bella		0;8.8	0;10.22	0;11.6		0;10.2	0;10.22			0;11.6	0;11.6	
Maebh		0;11.10										
Jennifer	0;10.0	0;6.16	0;10.27									
Morgan	0;8.2											

Table 6 Vocal Motor Schemes attained by all babies from whom data were collect in the present study, and the ages at which they were attained.

Reliability: coding for expression

Three research assistants, all students of linguistics blind to the hypothesis and research questions, performed reliability checks for coding for expression in a sample of 13 videos

(11.3% of the audio-visual corpus). Overall coder agreement was 82.3% (range 80-85.7%), Cohen’s kappa = 0.61. This proportion includes cases where the author of this thesis coded an expression for which the research assistant made no entry – hereafter ‘Ø-entries’. Where Ø-entries occurred, these tended to be facial expressions or (parts of) vocalisations that were either subtly articulated or of very short duration, often co-occurring with background noise e.g., brief smiles, grunts, hiccoughs, and singleton or whispered babbles. Overall coder agreement after excluding Ø-entries was 95.6% (range 95.2-96%), Cohen’s kappa = 0.92. Still-frames were not extracted from utterances or expressions that were coded as Ø-entries.

Reliability: coding for expression at a consistent point in development

One research assistant (GC) coded 5 videos of baseline visits (4.4% of the audio-visual corpus) to check the reliability of the author’s identification of non-babbles and smiles. Overall coder agreement including Ø-entries was 85.7% (range 63.6-94.1%), Cohen’s kappa = 0.7. Agreement for non-babbles was 88.24% (range 60-100%) and for smile/laugh, 81.8% (range 77.8-100%). After excluding Ø-entries, coder agreement rose to 96% (range 94.1-100%), Cohen’s kappa = 0.89. Agreement for non-babbles rose to 100% and for smile/laugh to 90% (range 83.3-100%).

Reliability: coding for expression within one baby at different times in development

A second research assistant (AP) coded a further 4 videos (3.5% of the audio-visual corpus) taken from one baby at four different home visits occurring between 0;5.1-0;8.26, again, to check the reliability of the author’s identification of vocalisations vs. smiles and laughs. Due to this research assistant’s non-specialist background with regard to pre-linguistic development, both babbles and non-babbles were collapsed into a single category: *vocalisation*. Overall coder agreement for this sub-sample was 80% (range 66.7-86.2%), Cohen’s kappa = 0.53. Coder agreement for *vocalisation* was 82.6% (range 66.7-90%) and for smile/laugh, 62.55% (range 50-100%). After excluding Ø-entries, overall coder agreement was 92.5% (range 84.6%-100%), Cohen’s kappa = 0.84. Coder agreement rose to 96.8% (range 84.6-100%) for vocalisations and 95.6% (range 83.3%-100%) for smiles/laugh.

Reliability: coding for expression across babies at different times relative to Vocal Motor Scheme attainment

A third research assistant (LE) coded a final 4 videos (3.5% of the audio-visual corpus), sampled across four babies at four different stages of development after the onset of babble. The sample comprised one video recorded immediately after the onset of babble at 0;7.5; one video in which a second baby attained their first Vocal Motor Scheme at 10;5.6; one video in which third a baby attained their second Vocal Motor Scheme at 0;8.6; and one video in which a fourth baby attained their third Vocal Motor Scheme at their final visit at 0;11.23. Some of these videos contained a small number of possible word forms as an additional category alongside babble, non-babble, and smiles. Discriminating babbles from possible early word forms without familiarity with the baby and without supplementary information from caregivers can be challenging and so a higher degree of non-agreement was expected in this category than in others. However, more non-agreement was instead seen in the non-babble category.

Overall inter-coder agreement for these videos was 56.1% (range 36.9-65.8%), Cohen's kappa 0.34. By category, coder agreement was 60.4% for babble, 30.8% for non-babble, 64.9% for smile/laugh, and 67.9% for possible word forms. Since agreement for babble and possible words was closely similar and since the video sample contained only 10 possible words, these two categories are collapsed hereafter. Excluding Ø-entries, coder agreement rose to 83.2% (range 70.8-87.8%), Cohen's kappa = 0.61. Coder agreement by category rose to 87.4% for babble and possible words, 60% for non-babble, 82.8% for smile/laugh.

It is possible that inter-coder discrepancies may have resulted from the fact that whilst LE is a student of linguistics, they had limited expertise in pre-linguistic phonological development. Alternatively, given the high proportion of Ø-entries in LE's coding, it is possible that LE's coding was less exhaustive than that of the author of this thesis. For these reasons, additional coding was carried out for these videos by another research assistant (KO), who is a paediatric speech and language therapist with 13 years of clinical phonetic transcription experience.

Reliability: transcription

This fourth research assistant (KO), blind to the intended use of transcribed babble utterances in the present study, transcribed babble and possible word forms in four randomly sampled three-minute-long excerpts from the same four videos coded by LE. McGillion et al.'s (2017) procedure for sampling and testing the reliability of babble and early word transcription used in Vocal Motor Scheme identification was adopted here. The Microsoft Excel (Microsoft, 2018) function =RAND()*30 was used to randomly generate a start time within 30 minutes for these three-minute excerpts. The first five minutes were avoided as babies were sometimes less voluble at the very beginning of a recording session. A randomly generated value of 00:09:40 was applied as the excerpt start time for each of the four videos. Overall coder agreement was 77.1%, Cohen's kappa = 0.65. Table 7 shows coder agreement by consonant. Most cases of disagreement centred around /ʔ/, /t,d/, and /k,g/. There were five Ø-entries where the author of this thesis transcribed a consonant and KO did not or vice versa. Three of these Ø-entries were cases where the author transcribed a glottal stop, but KO did not transcribe any speech sound e.g., [daʔdaʔda] and [dadada]. In another Ø-entry, the author transcribed [m̥a], whilst KO transcribed '*lip smack + vowel on release*'. In the final case, the author transcribed [gu], but KO made no entry. When glottal stops and the two remaining Ø-entries were removed, agreement rose to 85.71%, Cohen's kappa = 0.74.

	p/b	m	ɸ	t/d	n	ʒ	j	k/g	w	ʔ
total	3	3	1	7	1	1	15	11	2	4
% agreed	100%	100%	100%	57%	100%	0%	93.%	73%	100%	25%

Table 7 Inter-rater agreement between FO and KO by phoneme before the exclusion of glottal stops and Ø-entries.

Orofacial asymmetry analysis

The coded and transcribed videos were used to identify still-frames for orofacial asymmetry analysis. The precedent in orofacial asymmetry analysis is to extract and analyse the still-frame in which the mouth is open the widest during a vocalisation or facial expression (e.g., Graves et al., 1982; Holowka & Petitto, 2002a). This still-frame was identified using the frame-by-frame viewing function in ELAN version 4.9.4 (2016) In naturalistic data, participants are free to move around, which can impede data collection. As such, additional criteria previously used in naturalistic orofacial asymmetry studies with primates were used here to identify suitable still-frames for analysis (Hook-Costigan & Rogers, 1998; Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). Suitable still-frames must i) be in focus and ii) show the baby facing the camera, judged by whether both eyes, ears, and sides of the mouth were visible. Additional criteria were added to tailor the methodology to work with babies: i) the lighting must permit identification of the facial features; and ii) where the mouth was partially or wholly blurred or obscured during the actual point of maximal gestural articulation (e.g., if the baby moved suddenly or obscured the mouth with the hands or a toy) the still-frame closest to this point, before or after the actual peak in lip opening, during which the mouth was unobscured and in focus was extracted instead. Up to 10 still-frames were extracted per category, per home visit for each baby, giving a total sample of 1501 analysable still-frames: 585 of non-babble, 540 of smile/laugh, and 376 of babble. Table 8 shows the distribution of these still-frames across babies and categories. Babble still-frames were fewer in number because 14/16 baseline home visits (two baseline visits for each of the eight babies) contained no babble but did contain non-babble vocalisations and smiles and laughs.

	Freya	Benji	Cameron	Orelia	Fred	Adelaide	Arthur	Leif	TOTAL
non-babble	98	88	81	95	50	56	42	75	585
smile/laugh	99	89	72	91	47	40	37	65	540
babble	57	61	35	69	33	36	28	57	376
TOTAL	254	238	188	255	130	132	107	197	1501

Table 8 Distribution of still-frames extracted for analysis across babies and categories.

Analysis was carried out in Adobe Photoshop Elements 13 (Adobe Systems Incorporated, San Jose, CA) following the methods used in primate studies (Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). First, images were enlarged using the *Zoom* function and inspected for quality and clarity. In darker images, brightness and contrast were manipulated for ease of viewing. In a minority of more pixelated images, the *Despeckle* function was used to smooth the image to allow faster and easier identification of the boundaries of the facial features. If the facial features remained difficult to distinguish after these manipulations, then the still-frame was rejected. Next, head position was checked using the *Grid* function to ascertain whether the baby's ears, nostrils, and medial and lateral canthi (inner and outer eye corners) were level. If necessary, the image was rotated to level using the *Free rotate* function. Still-frames were then cropped to show only the baby's face. The baseline visits before babble emergence facilitated the identification of minor individual asymmetries in facial and cranial physiology like slight left vs. right differences in ear height, eye size, or nostril width. For example, Benji's right eye was slightly narrower and higher than his left eye, Freya and Cameron both had slightly larger and more protruding left than right ears, and Cameron tended to open the right side of his mouth more widely in both smiles and non-babble vocalisations (see Figure 6 below). No baby in the sample exhibited large enough left vs. right physiological differences to impede analysis.

Figure 6 Left: Benji showing subtle difference in the size and shape of the eyes. Middle: Freya showing larger and more protruding left ear. Right: Cameron showing larger and more protruding left ear and rightward asymmetry in the shape of the lip opening.

Using the ‘Line’ tool, horizontal lines were drawn between the lateral canthi (Line 1, orange) and the medial canthi (Line 2, green, see Figure 7). These lines were used to find the distance in pixels between the lateral canthus and the medial canthus for each eye and the resulting distance values were used to calculate a Facial Asymmetry Index between 1 and -1 for the baby’s eyes (eFAI) using the standard formula for calculating laterality indices (Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012):

$$FAI = (R - L) \ / \ (R + L)$$

Figure 7 Image of Fred with lateral canthi line (Line 1, orange), medial canthi line (line 2, green) and ala line (Line 3, blue).

These eFAIs indicate whether the baby is facing the camera directly or is minutely turned away to the left or right. Still-frames with an eFAI ≥0.6 were rejected.

These canthal lines are also used to identify the midline of the mouth in the primate studies mentioned earlier (Hook-Costigan & Rogers, 1998; Fernandez Carriba et al., 2002a, 2002b; Losin et al, 2008; Wallez et al., 2012; Wallez & Vauclair, 2012). An orofacial asymmetry study with babies by Schuetze and Reid (2005) instead used the philtrum to find the midline of the mouth. The present study adapts these pre-existing methods by using both canthal and philtral midlines. However, whilst Schuetze and Reid's method relied on visual inspection to identify the centre of the philtral groove, in the present study, the nose was used to identify this point. This adaptation was piloted in Oxley et al. (2014) and found to achieve more reliably accurate identification the centre of the philtral grove than visual inspection of the philtrum alone. A third line (Line 3, blue, see Figure 7) was drawn between either the outer edges of the nasal vestibule (nostril hole) or the outer edges of the ala (outer nostril skin) depending on which of these parts of the nose was most clearly visible. The midpoints of Lines 1-3 were identified, and perpendicular lines were drawn from each line to bisect the mouth into two *hemimouths* (see Figure 8). Still-frames in which the midlines i) wholly or partially overlapped and ii) bisected the philtral groove were deemed optimal. These represented the great majority of cases. Still-frames in which more than one of these midlines lay outside the philtral groove or in which midlines were separated by a gap of 1-3 pixels (dependent on the total number of pixels contained in the cropped image) were rejected.

The most accurate midline (as determined by overlap and position within the philtrum) was selected and the area of each hemimouth either side of this midline was identified by using the *Quick select* tool to trace around the inner edge of the lip opening (see Figure 8). In previous studies, the *Freehand* tool has been used to trace the inner perimeter of the lip opening (Wallez & Vauclair, 2012). The *Quick select* tool was used here to expediate hemimouth measurement and minor adjustments were made using the *Freehand* tool only where the *Quick select* tool over- or under-selected pixels that formed parts of the lips, tongue, teeth, and gums. Hemimouth area was measured in pixels using the *Histogram* function and these values were used to create an Oral Asymmetry Index (OAI) for each still-frame using the same standard formula for calculating laterality indices:

$$(R - L) \ / \ (R + L).$$

Positive OAIs indicate greater right-sided mouth opening, implying left hemisphere lateralisation, whilst negative OAIs indicate greater left-sided mouth opening, suggesting right hemisphere lateralisation. Following Losin et al. (2008) and Wallez et al. (2012), these OAIs were adjusted by subtracting the eFAI from the OAI for each still-frame. The resulting values were then subjected to statistical analysis.

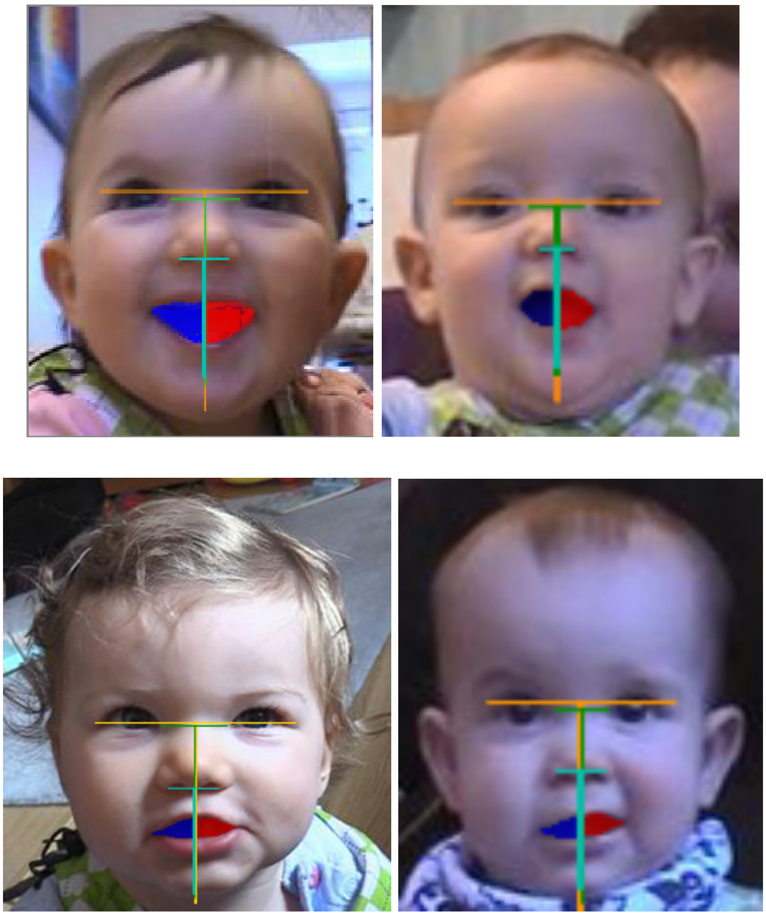


Figure 8 Top left to bottom right: images of Orelia, Arthur, Adelaide, and Leif with lines 1-3 and their midlines. The hemimouths are traced and highlighted. The right hemimouth is highlighted in blue and the left in red.

Reliability

A final research assistant (JKH) was trained in this protocol for measuring laterality via hemimouth area measurement, and calculated OAIs for a quasi-random sample of 152 still-frames from all eight babies, representing a range of points across the data collection period for each baby (10.12% of the total OAI corpus). This research assistant was kept blind to the category of expression of each still-frame under analysis and to the significance of orofacial asymmetry to the research questions explored in the present study. This sub-sample included 52 still-frames of non-babble, 52 of smile/laugh, and 48 of babble, selected randomly across each baby and home visit. Following Losin et al. (2008), a Pearson product moment correlation co-efficient was calculated, and this indicated strong agreement (r = 0.63 , df = 150, $p < 0.001$). Following Losin et al. (2008) an intra-class correlation co-efficient (ICC) was also calculated using a two-way mixed effects model for absolute agreement of single laterality measurements. The package *psych* (Revelle, 2022) in R Studio (R Development Core Team, 2021) was used to calculate the ICC. This fell within the 'good' range at 0.76 (Koo & Li, 2016). The results of this process are detailed in Table 10 below, in line with Koo and Li's recommendations for reporting ICCs.

baby	Freya	Benji	Cameron	Orelia	Fred	Adelaide	Arthur	Leif	TOTAL
non-babble	5	7	7	6	6	6	7	8	52
smile/laugh	7	5	6	6	7	6	7	8	52
babble	6	6	5	6	5	6	6	8	48
TOTAL	18	18	18	18	18	18	20	24	152

Table 9 Number and distribution of still-frames analysed by JKH.

	Intra-class correlation	Lower bound *(0.95 CI)*	Upper bound *(0.95 CI)*	*F* value	*df*	*P* value
Single measures	0.76	0.67	0.83	4.2	151	< 0.001

Table 10 Results of ICC calculation in R Studio, using a two-way mixed effects model for absolute agreement of single laterality measurements.

Coding for babble subtype and sequentiality

Following OAI generation, the timestamp and visit date of each still-frame was cross-referenced with the corresponding video and transcript in order to code each still-frame for the sub-type of babble (marginal, singleton, reduplicated, or variegated) from which it was extracted, and the number of syllables produced in that utterance (mono, bi, or poly). These indicators of utterance complexity were used to test for effects of increasing articulatory skill and sequentiality on the laterality of babble. Performing this coding after still-frame analysis prevented implicit researcher bias. Respecting the primacy of consonants in identifying and transcribing babble, babble sub-type was determined by the consonants produced within an utterance. Babble containing one single initial, medial or final consonant were coded as singleton babble e.g., [da], [ib], or [ɛnɛ]. Babble utterances containing a single repeated initial or final consonant were coded as reduplicated babble e.g., [vev], [dɪdɪd], or [bɒbɒbɒbɒ]. Babble utterances containing more than one consonant with more than one place of articulation were coded as variegated babble e.g., [bɪð], [vʊdi], or [utigəʃɛw]. Non-canonical syllables produced before the emergence of adult-like syllables were coded as marginal babble. Because the relationship between babble type and utterance length and complexity is not straightforward, syllable count was added as an extra measure to examine the length and sequentiality of babble utterances. Tables 4.4 and 4.5 below show the number still-frames extracted from babble of each sub-type and syllable count produced by each child in this sample.

	Freya	Benji	Cameron	Orelia	Fred	Adelaide	Arthur	Leif	TOTAL
marginal	0	0	0	0	0	0	5	2	7
singleton	21	13	12	42	12	15	9	9	133
reduplicated	7	24	7	9	10	6	2	15	80
variegated	29	24	16	18	11	15	12	31	156
TOTAL	57	61	35	69	33	36	28	57	376

Table 11Number of babble utterances produced by each child, arranged by babble sub-type.

	Freya	Benji	Cameron	Orelia	Fred	Adelaide	Arthur	Leif	TOTAL
mono	10	7	8	28	4	9	4	8	78
bi	23	13	8	21	13	11	11	12	112
poly	24	41	19	20	16	16	13	37	186
TOTAL	57	61	35	69	33	36	28	57	376

Table 12 Number of babble utterances produced by each child, arranged by utterance length in syllables.

Reliability

Transcripts produced by KO were also used to test the reliability of the author's coding for babble subtype and sequentiality. Overall coder agreement for coding babble subtype was 77.78% (range 62.5-100%), Cohen's kappa = 0.65. Overall coder agreement for utterance length was 77.78% (range 57.14-100%), Cohen's kappa = 0.69. Table 13 shows coder agreement by babble subtype and babble utterance length.

babble sub-type	%	syllable count	%
singleton	62.5%	mono	57.14%
reduplicated	100%	bi	80%
variegated	87.5%	poly	100%

Table 13 Inter-rater agreement between FO and KO by subtype and utterance length.

Summary

In this chapter, three research questions arising from a consideration of the literature reviewed in Chapters 1-3 have been identified and the ways in which the present study explores these questions has been outlined. Potential contributions of the present study have been considered, and the rationale behind methodological decisions undertaken has been explained. The participants involved in the present study have been introduced and the data generated, and the methods used to generate and analyse it have been described. Measures undertaken to ensure reliability have also been explained. In the following chapter, the data generated using these methods are subjected to statistical analyses using linear mixed effects models.

Chapter 5 Results

In this chapter, the statistical methods used to analyse the laterality data collected in this study are discussed. Mixed effects models are employed for their robustness and flexibility, the construction, testing, and application of six mixed effects models are outlined, and the findings of these models are presented.

Introduction

Linear mixed effects models were used to examine the laterality data for between category differences, between sex differences, and for evidence of change in the degree of asymmetry associated with increasing age, articulatory experience, and utterance complexity. Linear mixed effects modelling was chosen for its robustness in analysing large and unbalanced corpora of longitudinal data, which may be influenced by coefficients with both constant and variable values – that, is by both fixed and random predictors.

The dependent variable in the models outlined in this chapter was Oral Asymmetry Index (OAI). As described in Chapter 4, positive OAIs indicate left laterality, whilst negative OAIs indicate right laterality, and OAIs of 0 indicate equal hemispheric involvement. Repeated measures of laterality were taken from each baby for each category both within and across visits, meaning that measures of laterality were non-independent. It was anticipated that each instance of each baby's own behaviour might show more similar laterality to other instances of their own behaviour than to instances of other babies' behaviour. *baby_name* was therefore included in these models as a random effect. The fixed effects were those coefficients that were not estimated by these models, for which all possible values were considered, and which varied in consistent ways across participating babies and data points (Gelman & Hill, 2006; Starkweather, 2010; Monsalves et al., 2020). These predictors included still-frame category, baby's age at the time of production, articulatory experience, and utterance complexity. The values within each predictor are outlined where relevant in later sections of this chapter. Mixed effects modelling can differentiate the proportion of variance in a sample that is introduced by each predictor and each value within each predictor, and so can generate more precise estimates of the effect size of each value and predictor than can be generated by modelling only fixed effects (Monsalves et al., 2020).

The distribution of data points (OAIs) in this study was unbalanced across babies and categories. Firstly, babies who began to babble later were visited fewer times than early babblers. Secondly, non-babble and smile/laugh still-frames were collected from the time of the first home visit, usually two visits before babies began to babble.[32] These facts resulted in unbalanced possible quotas of still-frames per baby and per category. Thirdly, the total number of OAIs generated was also partly determined by the availability of analysable still-frames within video recordings from each visit. During visits when babies were quieter or smiled or laughed less, the actual quota of still-frames available for these categories was more limited. As a result, the OAI corpus contains missing values, which mixed effects modelling can accommodate.

R Studio (version 4.1.2, R Development Core Team, 2021) was used for the statistical analyses. Overall, six models were constructed to address the research questions outlined in Chapter 4. The first model tested for an effect of category. The second and third models tested for one-to-one interactions between category and age, and category and articulatory experience. The fourth model tested for effects of utterance complexity on the laterality of babble. The final two models tested for interactional effects of age and utterance complexity. In the fifth model, the effect of utterance length in syllables was explored, and the sixth model explored the effect of babble subtype (i.e., singleton, reduplicated, and variegated babble). The summarised output of each model is presented in the following pages. The full output of each model is included in Appendix III for reference. The package *lme4* (Bates et al., 2015) was used to construct the models and *lmerTest* (Kunzetsova et al., 2017) was used to run them. The package *MuMIn* (Barton, 2022) was used to assess the amount of variance explained by each model. In this package, the marginal R^2 indicates the amount of variance explained by the fixed predictor/s and the conditional R^2 indicates the amount of variance explained by all fixed and random predictors combined. The results are

---

[32] (with the exception of Arthur and Leif who, as mentioned in Chapter 4, began to produce marginal babble during their second home visit)

reported in line with Monsalves et al.'s (2020) *Logical Explanations & Visualizations of Estimates in Linear mixed models* (LEVEL) system. Visuals were created using the packages *sjPlot* (Lüdecke, 2021) *ggplot2* (Wickham, 2016), *Tidyverse* (Wickham et al., 2019), *hrbrthemes* (Rudis, 2020) and *viridis* (Garnier et al., 2021).

Model 1: modelling category

The first model was intended primarily to test for an effect of category on laterality to determine whether Holowka and Petitto's (2002a) findings could be replicated using the methodology outlined in Chapter 4. However, the explanatory power of other predictors was also explored. An initial additive model was constructed containing all possibly relevant predictors. Partial models were then constructed excluding each predictor in turn. These partial models were tested against the full model to identify the least complex model capable of explaining the variance in a dataset (Candarli, 2022). The initial additive model was constructed as follows and applied to a dataset containing all OAIs from all categories from all babies:

(5.1) *lmer (oai ~ category + age_days + vms + sex + (1 | baby_name)*

The fixed predictor *category* contained three possible values: *babble, non-babble,* and *smile/laugh*. Developmental time is measured by age in days (*age_days*) for accuracy and consistency across the eight participating babies, most of whom were born in different months. The predictor *vms* had three possible values: *before_vms, during_vms,* and *post_vms*. OAIs were coded *before_vms* if drawn from visits before the baby had attained a Vocal Motor Scheme; as *during_vms* if drawn from the visit in which the first Vocal Motor Scheme emerged; or *post_vms* if drawn from a visit following the attainment of the first Vocal Motor Scheme (see Table 5, Chapter 4 for the number of *before_vms, during_vms,* and *post_vms* analysed for each baby). Only four babies attained more than one Vocal Motor Scheme during the data collection period, resulting in too small a sample of *post_vms* data to incorporate separate values for attainment of the first and second Vocal Motor

Schemes in this model. The predictor *sex* denotes the baby's assigned sex at birth as reported by their caregivers and had two possible values: male and female. Predictors relating to utterance complexity (*babble_subtype* and *syllable_count*) were excluded from this model since these predictors were codes applied to babble OAIs only (n = 369), and so were unlikely to explain a significant proportion of the variance observed in the full corpus of n = 1494 OAIs. OAIs for the n = 7 marginal babbles produced by Leif (n = 2) and Arthur (n = 5) were excluded from the corpus for all analyses because they did not meet key criteria for identifying babble – namely, rhythmicity and articulatory precision – leaving n = 369/376 OAIs for babble. It was anticipated that marginal babble, being less stable and showing less adult-like rhythmic ability and weaker neuromuscular control, may be less strongly or consistently lateralised than canonical babble.

Four further models were constructed excluding *category* (5.2), *age_days* (5.3), *vms* (5.4) and sex (5.4) in turn:

(5.2) *lmer (oai ~ age_days + vms + sex + (1 | baby_name)*

(5.3) *lmer (oai ~ category + vms + sex + (1 | baby_name)*

(5.4) *lmer (oai ~ category + age_days + sex + (1 | baby_name)*

(5.5) *lmer (oai ~ category + age_days + vms + (1 | baby_name)*

The fit of each partial model was examined by running ANOVAs comparing each partial model against the full model. The resulting *Pr (>Chisq)*, Akaike's Information Criteria (AICs), and Bayesian Information Criteria (BICs) were used to identify the best fitting model (Monsalves et al., 2020; Candarli, 2022). *Pr (>Chisq)* values of < 0.05 indicate that the excluded parameter explains a significant proportion of variance and should therefore be retained in the model. AIC and BIC assess goodness of fit and predictive power by penalising statistical models containing numerous parameters, of which some parameters have low explanatory power, by increasing the error of the model (Wagenmakers & Farell, 2004; Vrieze, 2012). Unlike *Pr (Chisq)*, AIC and BIC are comparative statistics so there is no absolute AIC or BIC value that signifies best fit – rather, the best fitting model is selected by identifying the model with the smallest AIC and BIC values (Vrieze, 2014). BIC applies a higher penalty than AIC. Smaller AIC and BIC values indicate better model fit (Wagenmakers & Farell, 2004; Gelman & Hill, 2006). AIC and BIC are used in conjunction in the analyses reported in this chapter since using AIC alone can result in overfitting models and using BIC alone can result in underfitting models (Vrieze, 2012). The results of this exclusion process are summarised in Table 14.

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.1)		-2249.9	-2202.1
(5.2)	**0.0003242**	-2237.8	-2200.6
(5.3)	0.4049	**-2251.2**	**-2208.7**
(5.4)	0.3456	**-2251.7**	**-2214.6**
(5.5)	0.8279	**-2251.8**	**-2209.3**

Table 14 Results of ANOVAs comparing partial additive models to a full additive model containing all relevant predictors: category, age_days, and vms. Significant Pr (>Chisq) value is shown in bold. Lowest AIC and BIC values are shown in bold.

Excluding the predictor (or parameter) *category* from the model (5.2) resulted in a significantly poorer fitting model containing a higher proportion of unexplained variance compared to the full model (*Pr(>Chisq)* = 0.0003242, *AIC* = -2239.8, *BIC* = -2207.2). Explanatory power and goodness of fit are not significantly altered when either *age_days* (*Pr(>Chisq)* = 0.4049) (5.3), *vms* (*Pr(>Chisq)* = 0.3456) (5.4), or *sex* (*Pr(>Chisq)* = 0.8279) was excluded from the model. Partial models retaining *category* yield smaller AIC and BIC values than the model excluding *category*. AICs for partial models retaining *category* and excluding *age_days*, *vms*, and *sex* in turn are closely similar at -2251.2, -2251.7, and -2251.8 respectively. The BICs for models retaining *category* and excluding these other predictors are less closely similar at -2208.7, -2214.6, and -2209.3 respectively, though all BICS for these latter three models are smaller than those values for the full model and the model excluding *category*. The predictors *age_days* and *sex* explain more variance than *vms,* though the proportion of variance explained by each of these predictors is still non-significant. The best fitting additive model was found to be one retaining only category:

(5.5) *lmer (oai ~ category + (1 | baby name))*

Table 15 summarises the results of this model. The reference level (Intercept) is the value *babble* of the parameter *category*.

A significant effect of *category* is observed. The estimated OAI for *babble* shows significant left laterality with a category estimate of 0.026 (t = 3.821, df = 0.3821, p = <0.001). OAI estimates for the other categories of orofacial gesture are significantly less left lateralised, with category estimates of -0.003 for *non-babble* (t = -3.738, df = 0.0015, p = <0.001) and -0.009 for *smile/laugh* (t = -4.458, df = 0.0015 p = <0.001). Appendix III: Model 1 details the full output of this model including category estimates shown to all decimal places. Both *non-babble* and *smile/laugh* show slight right laterality, though this does not diverge as far from zero as the estimate for *babble*. A greater between-category difference is seen for *babble* vs. *smile/laugh* than for *babble* vs. *non-babble*. Whilst the difference between category estimates is statistically significant, there is high within-category variance and marked between-category overlap amongst individual OAIs, as illustrated by the violin plot in Figure 9.

Little of the variance in laterality in each category of orofacial gesture that is observed in the full OAI corpus (see Figure 9) can be attributed to the fixed or random predictors in this model. The ICC of 0.01 indicates that there is little correlation between each baby's OAIs. This may be a result of the significant between-category difference in OAI estimates. The marginal R^2 indicates that 1.4% of the observed variance in attributable to the fixed predictor *category* and the conditional R^2 indicates that *category* and *baby_name* combined account for 2%. Because this additive model pools OAIs collected at different points in developmental time, it is possible that the residual variance (98%) may arise from interactions between *category* and these excluded predictors. Alternatively, this residual variance may indicate that, before 12 months of age, babies do not yet possess sufficient neuromuscular control or sufficiently specialised and lateralised neural networks to produce consistently lateralised orofacial movements across multiple instances over real and developmental time. These possible interactions are investigated in the next section of this chapter.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
babble (Intercept)	0.026	0.007	0.01 – 0.04	3.821	**<0.001**
non-babble	-0.028	0.008	-0.04 – -0.01	-3.738	**<0.001**
smile/laugh	-0.034	0.008	-0.05 – -0.02	-4.458	**<0.001**
Random Effects					
σ^2	0.01				
$^{\tau}00$ baby_name	0.00				
ICC	0.01				
^{N} baby_name	8				
Observations	1494				
Marginal R^2	0.014				
Conditional R^2	0.020				

Table 15 Results of mixed effects model testing for an effect of category on laterality. Reference level is babble. Significant p-values are shown in bold.

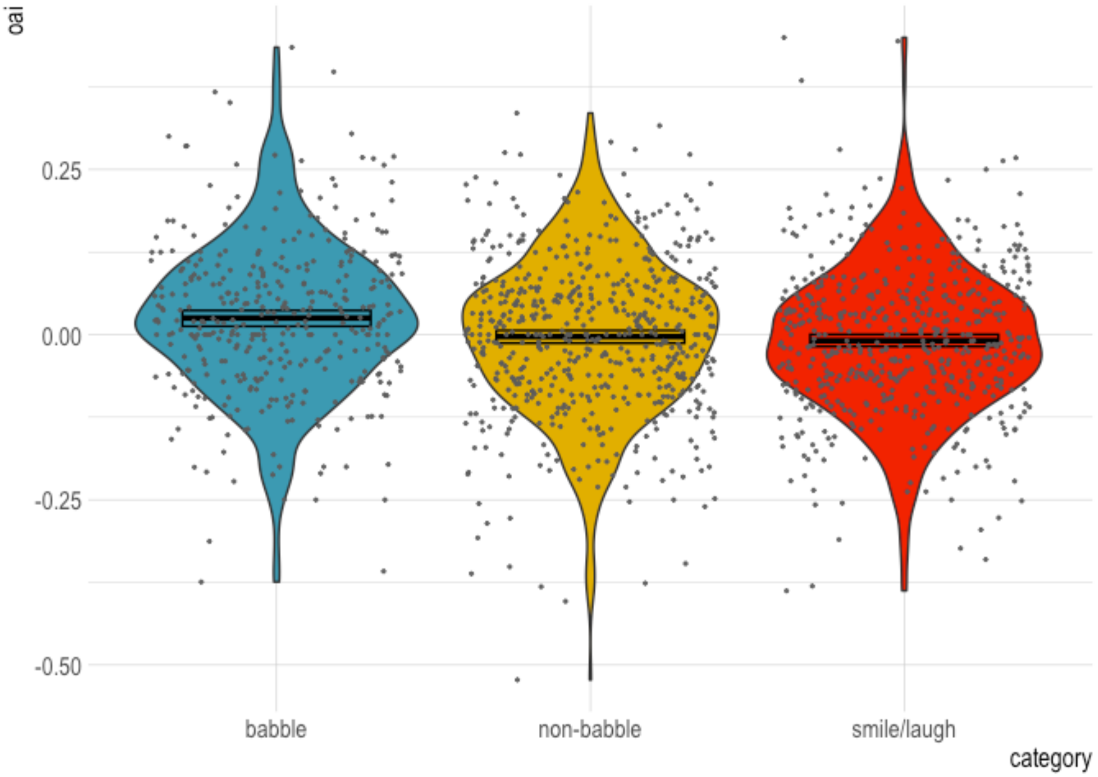


Figure 9 Violin plot showing distribution of OAIs by category and within-category variance. Bootstraps show category means and confidence intervals (0.95).

Models 2-3: modelling interactions between category and other predictors

Having found no significant effect of age and Vocal Motor Scheme attainment when laterality data from all categories were pooled together, two further models were constructed to test for interactions between category and each of these predictors. Running models with small subset sizes can yield unstable estimates and the OAI corpus contained insufficient data points to run a single model containing all one-to-one interactions between *category* and *age_days* and *vms*. Two separate models were therefore constructed to test for effects of these interactions.

Model 2: modelling category and age

One model was constructed to test for changes in laterality associated with the passage of developmental time. Research reviewed in Chapter 3 has found evidence that marked structural and functional changes take place in the parts of the brain involved in vocalisation and auditory processing during the first year of life (Witelson & Pallie, 1973; Chi et al., 1977; Thelen & Smith, 1994; Bogolepova & Malofeeva, 2001; Dehaene-Lambertz et al., 2002; Paterson et al., 2006; Mareschal et al., 2007; Dehaene-Lambertz et al., 2010; Perani et al., 2011; Mahmoudzadeh et al., 2013). Orofacial asymmetry and neuroimaging research with babies, toddlers, and children has identified changes in the laterality of emotional processing and emotional facial expressions persisting throughout infancy and childhood and perhaps into adolescence (Best & Queen, 1989; Rothbart et al., 1989; Schuetze & Reid, 2005; Workman et al., 2006; Nagy, 2012; Wallez & Vauclair, 2012; Lindell et al., 2017; Santamaria et al., 2020). An interactional model was constructed and applied to the full OAI corpus to investigate whether the degree of laterality in any category of orofacial gesture was associated with advances in age. This model was constructed as follows:

(5.6) *lmer (oai ~ category * age_days + (1 + category * age_days | baby name))*

The *category * age* interaction was initially included both as a fixed predictor and as random slope in order to create a more conservative model (5.6). It was anticipated that each baby's behaviour may vary to different degrees in each category of orofacial gesture at each point in developmental time. Including random slopes tests whether any significant effect of the fixed interaction that may be observed persists after the random variance (noise) that may be introduced by this interaction is 'cleaned'. This model was tested against a model excluding the random slope (5.7) and a model excluding the fixed interaction (5.8) in order to identify the better fitting model (Starkweather, 2010):

(5.7) *lmer (oai ~ category * age_days + (1 | baby name))*

(5.8) *lmer (oai ~ (1 + category * age_days | baby name))*

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.6)		*-2155.6*	*-2006.8*
(5.7)	1	**-2262.1**	**-2219.6**
(5.8)	**0.06803**	-2155.2	-2033.1

Table 16 Results of ANOVA comparing full category * age_days interactional model to a model excluding the random slopes (category * age_days) (5.7) and a model excluding the fixed category * age_days interaction (5.8).

The results of the ANOVA shown in Table 16 indicate the best fitting model to be model (5.7), which retains the fixed interaction *category* * *age_days* and excludes the random slopes (*AIC* = -2262.1, *BIC* = -2219.6). This fixed interaction explains a marginally significant proportion of the variance observed in this model and so removing it, as in model (5.8) results in a marginally significantly poorer fitting model (*Pr (>Chisq)* = 0.0683). Removing the random slope, however, as in model (5.8) does not significantly worsen the fit of the model (*Pr (>Chisq)* = 1), indicating that the random slopes do not explain a significant proportion of the variance.

Table 17 summarises the results of the best fitting model (5.7). The reference level (Intercept) is *babble*. Significant interactions are seen between the fixed predictors *category* and *age_days* for the values *babble* and *non-babble,* but not for *smile.*

When all predictors are at 0, the category estimate of -0.128 for babble is significantly different from 0 (*t* = -2.585, *df* = 0.001, *p* = 0.01) and indicates right laterality. With each one-unit increase in *age_days* (i.e., with each passing day, the OAI for *babble* is estimated to increase by +0.0004154 (see Appendix III: Model 2) and this increase passes the $p < 0.05$ threshold for significance (*t* = 3.208, *df* = 0.001, *p* = 0.001) indicating a gradual leftward shift in laterality over developmental time.

When *age_days* is at 0, the estimated OAI for non-babble is significantly different from that for babble at -0.0088 (*t* = 2.533, *df* = 0.001, *p* = 0.01) indicating left laterality. The estimated OAI for *non-babble* decreases by by -0.0004865 per day (*t* = -3.136, 0.001, *p* = 0.002) indicating a significant rightward shift in laterality over developmental time.

When *age_days* is at 0, the estimated OAI of -0.0353 for smile/laugh indicates right laterality and is not significantly different from the estimate for babble (*t* = 1.014, *df* = 0.002, *p* = 0.311) (see Appendix III: Model 2 for the full output of this model). The OAI for *smile/laugh* is estimated to decrease by -0.0002456 per day (*t* = -1.568, *df* = 0.002, *p* = 0.117), indicating rightward shift. This shift, however, is not significant.

The ICC of 0.01 indicates little intra-individual correlation in OAIs, which, again may indicate unstable laterality or may result from the significant between-category difference and the significant interaction between *age* and *category.* The marginal R^2 indicates that 2.7% of the variance observed in this model is attributable to the interaction between *category* and *age*, whilst the conditional R^2 indicates that the fixed and random predictors combined account for 3.2%.

The trajectories of actual OAIs over developmental time, and their variability in real time, are visualised in in Figure 10 below. Notably, these trajectories show evidence of non-linear change over time and of high variability in real time, particularly amongst babble OAIs.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
babble (Intercept)	-0.128	0.044	-0.20 – -0.03	-2.585	**0.010**
non-babble	0.119	0.047	0.03 – 0.21	2.533	**0.011**
smile/laugh	0.048	0.047	-0.04 – 0.14	1.014	0.311
age days * babble	0.000	0.000	0.00 – 0.00	3.208	**0.001**
age days * non-babble	-0.000	0.000	-0.00 – -0.00	-3.136	**0.002**
age days * smile laugh	-0.000	0.000	-0.00 – -0.00	-1.568	**0.117**
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby name	0.00				
ICC	0.01				
^{N}baby name	8				
Observations	1494				
Marginal R^2	0.027				
Conditional R^2	0.032				

Table 17 Results of mixed effects model testing for differential effect of age on the laterality of each category of orofacial gesture. Reference level is babble. Significant p-values are shown in bold.

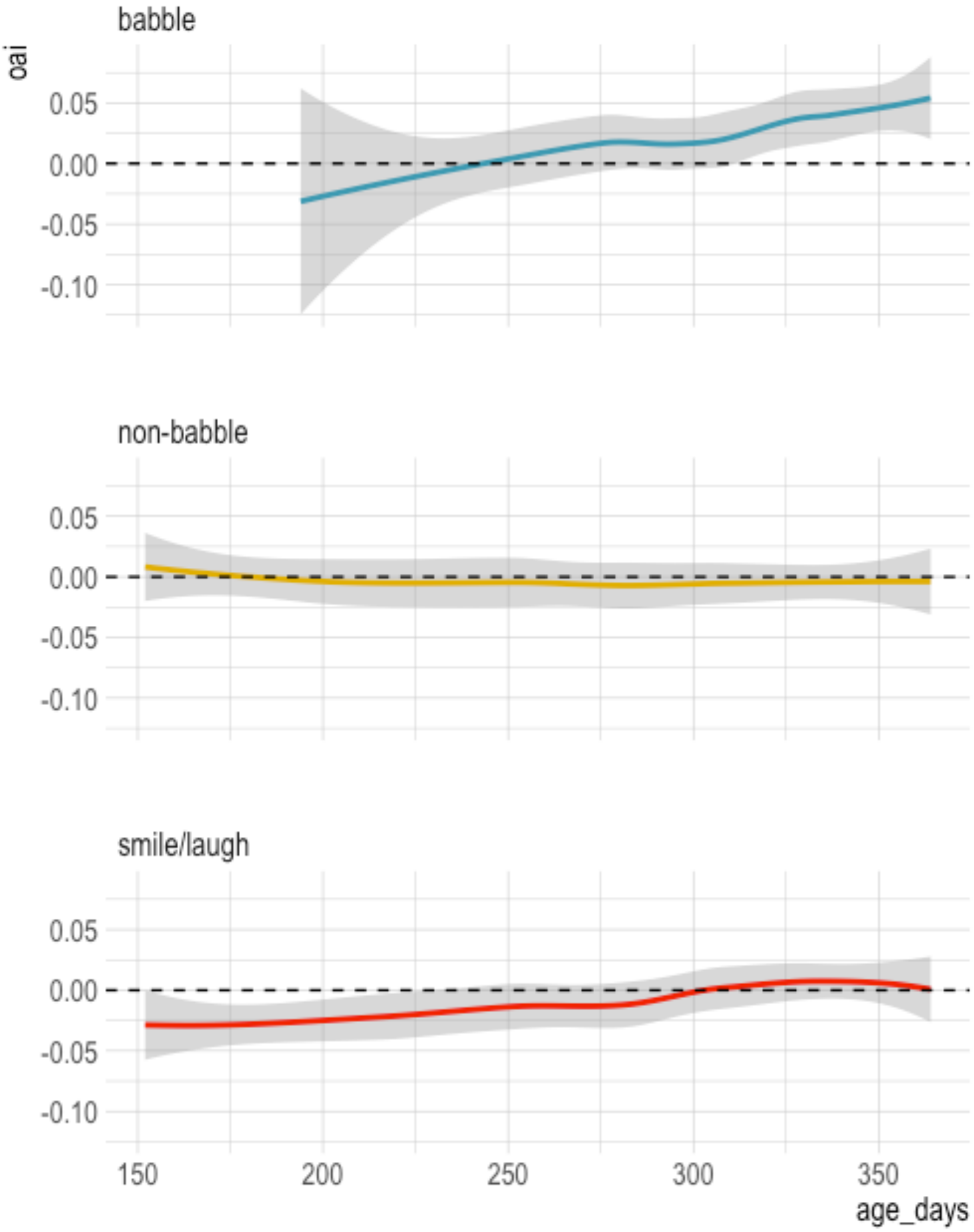


Figure 10 Spaghetti plot illustrating actual trajectories of OAIs for each category over time. Variability in real time is indicated by Shaded areas show real-time variability. Positive values indicate left laterality and vice versa.

Model 3: modelling category and articulatory experience

The next model was designed to examine whether any relationship could be observed between values within *category* and *vms*. A pilot to the present study found evidence to suggest that babble may become more left lateralised after babies attain a Vocal Motor Scheme (Oxley et al., 2014). In that pilot study, no such relationship was found between Vocal Motor Scheme attainment and the laterality or lateralisation of non-babble or smiles and laughs. This model was constructed as follows:

(5.9) *lmer (oai ~ category * vms + (1 + category * vms | baby name*))

This model was tested against an interactional model excluding the random slopes (5.10) and a model excluding the fixed interaction (5.11).

(5.10) *lmer (oai ~ category * vms + (1 | baby name*))

(5.11) *lmer (oai ~ (1 + category * vms | baby name*))

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.9)		*-2197.9*	*-1905.9*
(5.10)	0.7024	**-2247.3**	**-2188.9**
(5.11)	0.1205	**-2201.2**	**-1951.9**

Table 18 Results of ANOVA comparing full category * vms interactional model to a model excluding the random slopes (category * vms) (5.10) and a model excluding the fixed category * vms interaction (5.11). Lowest AIC and BIC values are shown in bold.

Table 18 indicates that the AIC and BIC were lower for the model excluding the random slopes (5.10) (*AIC* = -2247.3, *BIC* = -2188.9) than for model excluding the fixed interaction (5.11) (*AIC* = -2201.2, *BIC* = -1951.9). In both cases, values were lower than for the full model. Neither the fixed interaction nor the random slopes explained a significant proportion of the variance observed in this model (*Pr (>Chisq)* > 0.1). *before_vms, during_vms,* and *post_vms* OAIs were not available for each of the eight babies in the sample. Fred and Orelia were very voluble, and both babies attained one Vocal Motor Scheme during the first home visit after babble emergence at 0;9.12 and 0;7.2 respectively. Leif's first Vocal Motor Scheme emerged during the final home visit at 0;11.23. To ascertain whether this fact explained the non-significant *Pr (>Chisq)* value, the above process was repeated using a reduced dataset containing data from only those five babies for whom OAIs with all three values of *vms* were available: Freya, Benji, Cameron, Adelaide, and Arthur. The results of this exclusion process are shown in Table 19 below. Again, no significant variance in laterality in any category could be attributed to Vocal Motor Scheme attainment (*Pr (>Chisq)* > 0.1). Again, the model excluding the random slopes (5.10)‡ gave a lower AIC and BIC than the model excluding the fixed interaction (5.11)‡ and both models gave lower values than the full model.

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.9)‡		*-1324.5*	*-1059.5*
(5.10)‡	0.9969	**-1389.9**	**-1336.9**
(5.11)‡	0.3254	**-1331.3**	**-1104.8**

Table 19 Results of ANOVA comparing full category * vms interactional model to a model excluding the random slopes (category * vms) (5.10)‡ and a model excluding the fixed category * vms interaction (5.11)‡, using ‡reduced dataset.

Since the exclusion process yielded similar results for both the full and reduced data sets, the best fitting model (5.10) was run using the full data set to minimise the risk of returning unstable estimates. The output of this model is summarised in Table 20 below.

At the time of Vocal Motor Scheme attainment (*during_vms*), the category estimate of 0.005 for babble indicates slight left laterality, but this estimate is not significantly different from 0 ($t = 0.332$, $df = 0.0489$, $p = 0.740$). The *before_vms* estimate for babble is higher than the *during_vms* estimate at 0.016 , indicating stronger left laterality before Vocal Motor Scheme attainment although this difference is not significant ($t = 0.568$, $df = 0.0015$, $p = 0.570$ (see Appendix III: Model 3 for full output).). The *post_vms* estimate is also higher than the *during_vms* estimate at 0.037. This effect, whilst marginally significant, does not pass the $p < 0.05$ threshold for significance ($t = 1.766$, $df = 0.0015$, $p = 0.078$).

Notably, the OAI estimates for babble remain above zero, indicating left laterality before, at the time of, and following Vocal Motor Scheme attainment. Model 2 revealed babble to be right lateralised around the time of emergence and to become left lateralised gradually over developmental time. Model 3 indicates that this leftward shift in laterality is not associated with advances in articulatory skill involved in developing a motor preference for producing a particular consonant because babble was already left lateralised before the first Vocal Motor Scheme emerged.

No significant interaction was seen between Vocal Motor Scheme attainment for *non-babble* or *smile/laugh*. At the time of Vocal Motor Scheme emergence, the estimated OAIs for these categories of orofacial gesture are not significantly different from the estimate for babble at -0.011 ($t = -0.749$, $df = 0.0015$, $p > 0.1$) and 0.002 respectively ($t = -0.179$, $df = 0.0015$, $p > 0.1$). The OAI estimates for *non-babble* both prior to and following Vocal Scheme emergence are lower than -0.011 by -0.007915 ($t = -0.319$, $df = 0.0015$, $p > 0.1$) and -0.01387 respectively ($t = -0.566$, $df = 0.0015$, $p > 0.1$), showing a non-linear pattern of change in laterality. Estimates for *smile/laugh* show a similar pattern. The estimated OAI for *before_vms* is -0.03183 lower than the estimate for *vms[at]* ($t = -1.282$, $df = 0.0015$ $p > 0.1$) and the estimate for *post_vms* is -0.01387 lower ($t = -1.257$, $df = 0.0015$, $p > 0.1$). The non-linear patterns of shift seen here for all categories of orofacial gesture may reflect the fact that the subsets created by *vms[at]*, being drawn from a single home visit per baby, were much smaller than those created by *before_vms* and *post_vms* and so were more vulnerable to inter- and intra-individual variability.

The ICC of 0.01 indicates little intra-individual correlation in OAIs. The marginal R^2 indicates that 2.2% of the variance observed in this model is attributable to the interaction between *category* and *age*, whilst the conditional R^2 indicates that the fixed and random predictors combined account for 3.2%. Figure 11 below illustrates the extent of the overlap in actual OAIs for each category before, at, and after the emergence of the first Vocal Motor Scheme.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
during_vms * babble (Intercept)	0.005	0.016	-0.03 – 0.04	0.332	0.740
during_vms * non-babble	-0.016	0.022	-0.06 – 0.03	-0.749	0.454
during_vms * smile/laugh	-0.003	0.022	-0.05 – 0.04	-0.179	0.858
before_vms * babble	0.011	0.019	-0.03 – 0.05	0.568	0.570
before_vms * non-babble	-0.008	0.025	-0.06 – 0.04	-0.319	0.750
before_vms * smile/laugh	-0.032	0.025	-0.08 – 0.02	-1.282	0.200
post_vms * babble	0.032	0.018	-0.00 – 0.07	1.766	**0.078**
post_vms * non-babble	-0.014	0.025	-0.06 – 0.03	-0.566	0.571
post_vms * smile/laugh	-0.031	0.024	-0.08 – 0.02	-1.257	0.209
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby name	0.00				
ICC	0.01				
^{N}baby name	8				
Observations	1494				
Marginal R^2	0.022				
Conditional R^2	0.032				

Table 20 Results of mixed effects model testing for interaction between Vocal Motor Scheme attainment and category. Reference level babble during the session when the first Vocal Motor Scheme emerged. Marginally significant p-value is shown in bold.

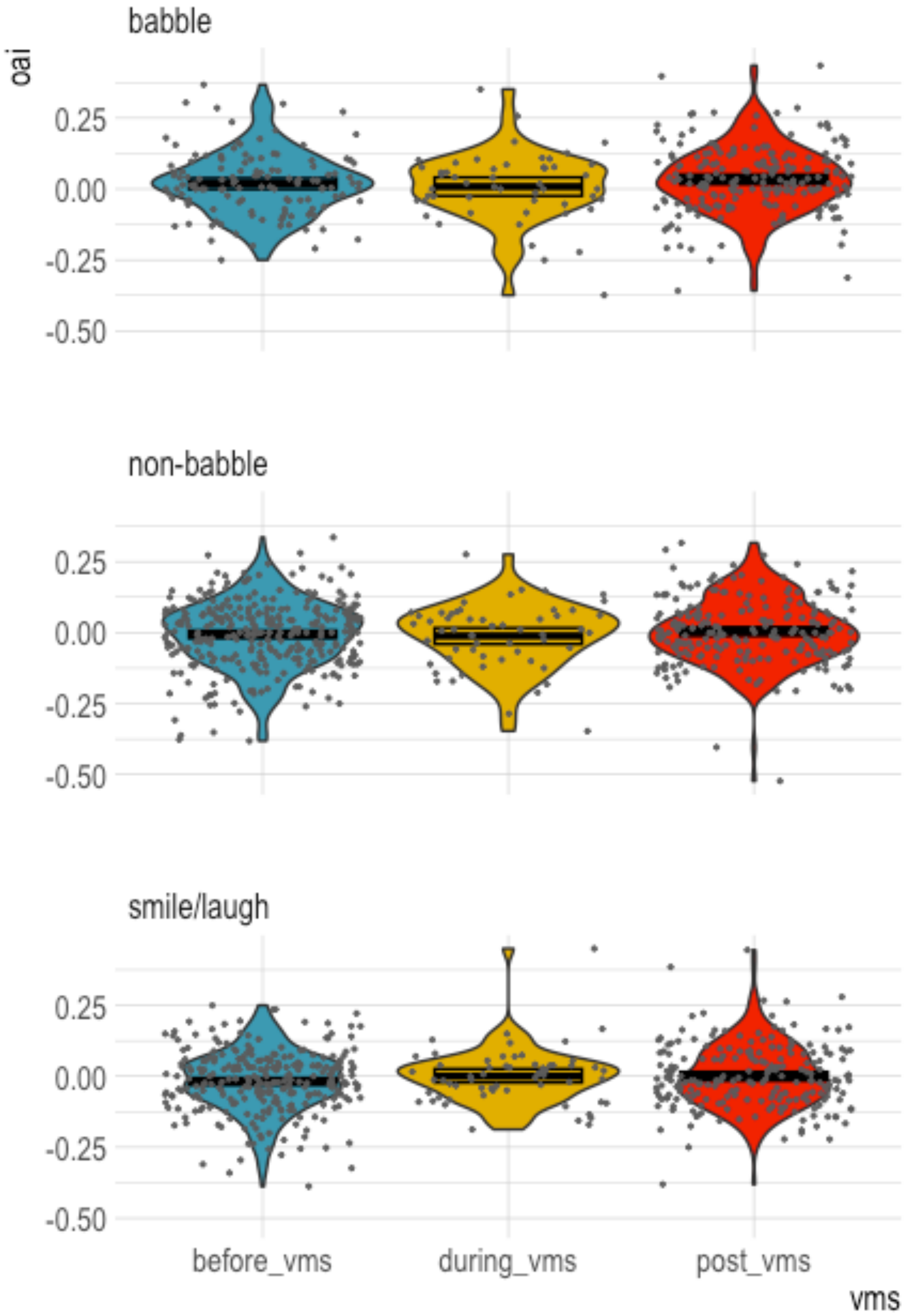


Figure 11 Violin plots showing variation in OAIs for each category as a function of Vocal Motor Scheme attainment and within-group variance. Bootstraps show category means and confidence intervals (0.95).

Model 4: modelling category and sex

This model was designed to investigate whether any sex differences in laterality exist between the female and male babies involved in the present study. Research in humans and non-human animals has found evidence of sex differences in laterality and temporal changes in laterality associated with hormonal differences between sexes (Wolf & Goodale, 1987; Hausmann et al., 1998; Frost et al., 1999; Fernández et al., 2003; Sommer et al., 2004; Clements et al., 2006; Lindell, 2006; Clements et al., 2006; Kaiser et al., 2007; Wallentin, 2018). An initial model was constructed as follows (5.12):

(5.12) lmer(oai ~ category * sex + (1 + category * sex | baby_name), data = OAI_data)

This model was tested against an interactional model excluding the random slopes (5.13) and a model excluding the fixed interaction (5.14)

(5.13) lmer(oai ~ category * sex + (1 | baby_name), data = OAI_data)

(5.14) lmer(oai ~ (1 + category * sex | baby_name), data = OAI_data)

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.12)		*-2219.4*	*-2070.7*
(5.13)	0.7778	**-2244.4**	**-2201.9**
(5.14)	**0.07319**	-2219.3	-2097.2

Table 21 Results of ANOVA comparing full category * sex interactional model to a model excluding the random slope (category * age_days) (5.7) and a model excluding the fixed category * age_days interaction (5.8).

The results of the ANOVA shown in Table 21 indicate the best fitting model to be model (5.13) which retains the fixed interaction *category * sex* and excludes the random slopes (-*AIC* = -2244.4, *BIC* = -2201.9). This fixed interaction explains a marginally significant proportion of the variance observed in this model and so removing it, as in model (5.14) results in a poorer fitting model (*Pr (>Chisq)* = 0.07319). Removing the random slope, however, as in model (5.13) does not significantly or marginally significantly worsen the fit of the model (*Pr (>Chisq)* = 0.778), indicating that the random slopes do not explain a significant proportion of the variance. Table 22 summarises the results of the best fitting model. The reference level is the interaction between femaleness and babble.

An effect of sex is observed for non-babble and smile/laugh but not for babble. The estimate of 0.027 for babble in female babies indicates significant left laterality (*t* = 2.582*, df* = 0.1844, *p* = 0.018). This estimate of 0.031 for babble in male babies is not significantly different (*t* = -0.260*, df* = 0.2184, *p* = 0.798).

The estimate of 0.001 for non-babble in female babies, indicates significantly weaker right laterality (*t* = 0.001*, df* = 0.0015, *p* = 0.016) than the estimate for babble. However, in male babies, the estimated OAI of 0.028 is not significantly different from the estimated OAI for babble in girls (*t* = -0.077*, df* = 0.0015, *p* = 0.939). This model compares all interactions to the reference level *female * babble* and so does not report one-to-one comparisons of laterality for each category. However, the fact that *non-babble* in male babies shows *stronger* left laterality than *babble* in female babies indicates indirectly that non-babble in male babies shows significantly different laterality in the male babies compared with the female babies in the present study. The difference in the laterality of babble (OAI = 0.027) and non-babble (OAI = 0.001) was significant for female babies, meaning that the difference in the laterality of non-babble in male babies (OAI = 0.028) and of non-babble in female babies (OAI = 0.001) would also be significant.

The estimated OAI for smile/laugh in female babies is -0.013, indicating right laterality and a significant difference in the laterality of babbles compared to smiles in female babies (t = -*3.441*, df = 0.0015, p = 0.000). The estimated OAI of 0.037 for smile/laugh in male babies in, once again, not significantly different from the estimate of 0.027 for babble in female babies. This, again, suggests indirectly that male and female babies show significantly patterns of laterality for smile/laugh.

The ICC of 0.01 indicates little intra-individual correlation in OAIs. The marginal R^2 indicates that 1.5% of the variance observed in this model can be attributed to the fixed interaction between category and sex, and the conditional R^2 indicates that the fixed and random predictors combined account for 2.3%.

Actual OAIs for male and female babies for each category are visualised in Figure 12. This figure shows a close relationship between male and female babies' OAIs for all categories.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
female * babble (Intercept)†	0.027	0.011	0.01 – 0.05	2.582	**0.018**
female * non-babble	-0.026	0.011	-0.05 – -0.01	-2.400	**0.017**
female * smile/laugh	-0.040	0.012	-0.06 – -0.02	-3.441	**0.000**
male * babble	0.004	0.014	-0.03 – 0.02	-0.260	0.798
male *non-babble	0.001	0.015	-0.03 – 0.03	-0.077	0.939
male * smile/laugh	0.010	0.016	-0.02 – 0.04	0.673	0.501
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby name	0.00				
ICC	0.01				
^{N}baby name	8				
Observations	1494				
Marginal R^2	0.015				
Conditional R^2	0.023				

Table 22 Results of mixed effects model testing for differential effect of sex on the laterality of each category of orofacial gesture. Reference level female * babble. Significant p-values are shown in bold.

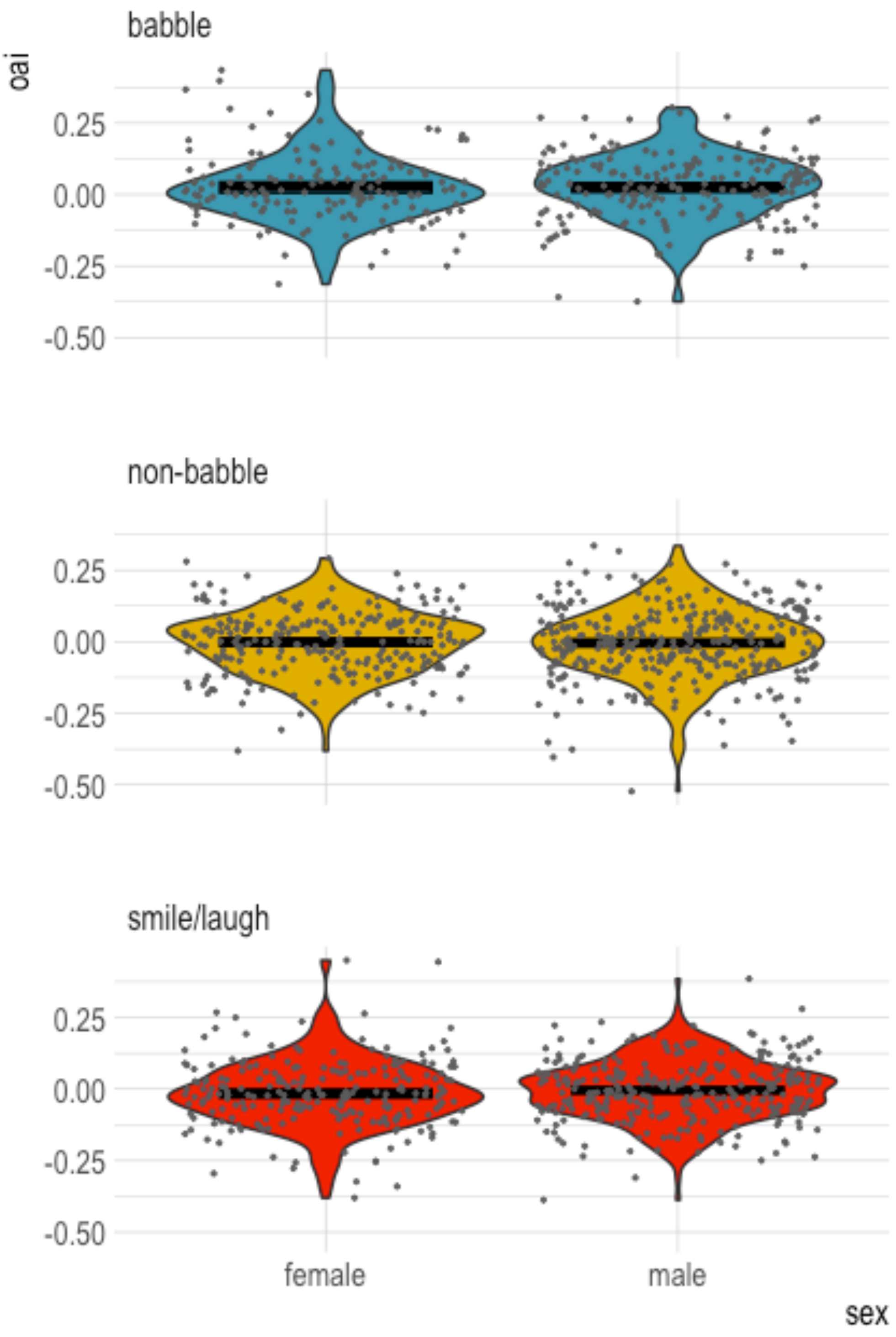


Figure 12 Violin plots showing variation in OAIs for each category for male and femal babies and within-group variance. Bootstraps show category means and confidence intervals (0.95).

Model 5: modelling utterance complexity

Having established that developmental time effects the laterality of babble significantly and articulatory experience does so near-significantly, a further model was constructed to test for effects of other indicators of advancing articulatory experience. It seems likely that the gradual leftward shift in the laterality of babble demonstrated by Model 2 may be at least partly associated with experience-dependent cerebral reorganisation of the type discussed in Chapter 3 (Van Der Knaap & Valk, 1990; Hayakawa et al., 1991; Thelen & Smith, 1994; Johnson, 2000; Bogolepova & Malofeeva, 2001; Cheour et al., 2002; Dehaene-Lambertz et al., 2002; Paterson et al., 2006; Mareschal et al., 2007; Su et al., 2008; Dehaene-Lambertz et al., 2010; Minagawa-Kawai et al., 2011; Petanjek et al., 2011; Mahmoudzadeh et al., 2013; Paredes et al., 2016; Williamson & Lyons, 2018; Yang et al., 2019; Grotheer et al., 2022). Some laterality research has found stronger asymmetries during the production and perception of longer or more complex sequences of oral (and manual) gestures (Wolf & Goodale, 1987; Salmelin & Sams, 2002; Saarinen et al., 2005; Hodgson et al., 2021). In particular, Ramsay (1980, 1984) has suggested that the emergence of reduplicated and variegated babble may be associated with some degree of hemispheric specialisation.

In the following models, two types of utterance complexity - utterance length in syllables and babble subtype – were used as indicators of articulatory skill. The predictor *syllable_count* contained three possible values: *mono, bi,* and *poly*. The predictor *babble_subtype* contained a further three possible values: *singleton, reduplicated,* and *variegated*. Since these predictors can only meaningfully be applied to babble utterances, these models were run using a reduced corpus containing only laterality data for babble. OAIs for the seven marginal babbles produced by Leif (n = 2) and Arthur (n = 5) were excluded from the sub-corpus for all analyses because they comprised too small a subset to give a stable category estimate.

An initial additive model was constructed containing both predictors as follows:

(5.15) *lmer (oai ~ syllable_count + babble_subtype + (1 | baby_name)*

Partial models were constructed, excluding *syllable_count* (5.13) and *subtype* (5.14) in turn, and these were tested against the full model to identify the best fitting model (see table 5.8 below).

(5.16) *lmer (oai ~ babble_subtype + (1 | baby_name)*

(5.17) *lmer (oai ~ syllable_count + (1 | baby_name)*

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.15)		*-539.28*	*-511.91*
(5.16)	**0.0839**	-538.33	-518.77
(5.17)	0.1782	**-539.83**	**-520.28**

Table 23 Results of ANOVAs comparing partial additive models to a full additive model containing all relevant predictors: syllable_count (5.13) and babble_subtype (5.14). Significant Pr (>Chisq) value is shown in bold. Lowest AIC and BIC values are shown in bold

Table 23 indicates the best fitting model was indicated to be one retaining syllable count but excluding *babble_subtype* (5.17). *Syllable_count* explains a marginally significant proportion of the variance in the laterality of babble (*Pr (>Chisq)* < 0.1), while *babble_subtype* does not (*Pr (>Chisq)* > 0.1). The AIC and BIC for this model (5.17) are in agreement with one another -539.83 and -520.28 respectively.

The results of this best fitting model are presented in Table 24 below (see Appendix III: Model 5 for full output). The reference level (Intercept) is bisyllabic babble. No significant differences are observed between the estimated OAI of 0.018 (*t* = 1.364, *df* = 0.2492, *p* =

0.185) for bisyllabic babble utterances and the estimates of 0.019 ($t = 0.035$, $df = 0.0366$, $p > 0.1$) for monosyllabic babble and 0.034 ($t = 1.189$, $df = 0.366$ $p > 0.1$) for polysyllabic babble. While the estimates for babble of one or two syllables are closely similar, the estimate of 0.034 for babble utterances composed of three of more syllables is markedly higher. In adults, longer and more complex utterances also show more strongly left lateralised orofacial movement and neural activity (Wolf & Goodale, 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). Actual OAIs for monosyllabic, bisyllabic, and polysyllabic babbles are visualised in Figure 13 for reference. OAIs for monosyllabic babbles are fewer in number and show wider dispersion than the other categories, though high variability is seen in all categories.

The ICC of 0.03 is higher than that of previous models in this chapter but still low. The marginal R^2 indicates that 0.5% of the variance observed is explained by the fixed predictor *syllable_count*, while the fixed and random predictors combined account for 3.2%. Since this model examines laterality data from babble only, and since no significant effects of any predictor are observed, it is likely that this residual variance (96.8%) is attributable to unstable orofacial asymmetry across instances of production. However, it should be noted that this model pooled laterality data across ages. For this reason, two further models were constructed to test for interactional effects between age and utterance complexity. These models are outlined in the following sections of this chapter.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
bi (Intercept)	0.018	0.013	-0.01 – 0.04	1.364	0.173
mono	0.001	0.017	-0.03 – 0.03	0.035	0.972
poly	0.016	0.014	-0.01 – 0.04	1.189	0.235
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby_name	0.00				
ICC	0.03				
^{N} baby_name	8				
Observations	369				
Marginal R^2	0.005				
Conditional R^2	0.032				

Table 24 Results of mixed effects model testing for an effect of utterance length on the laterality of babble. Reference level is bisyllabic babble.

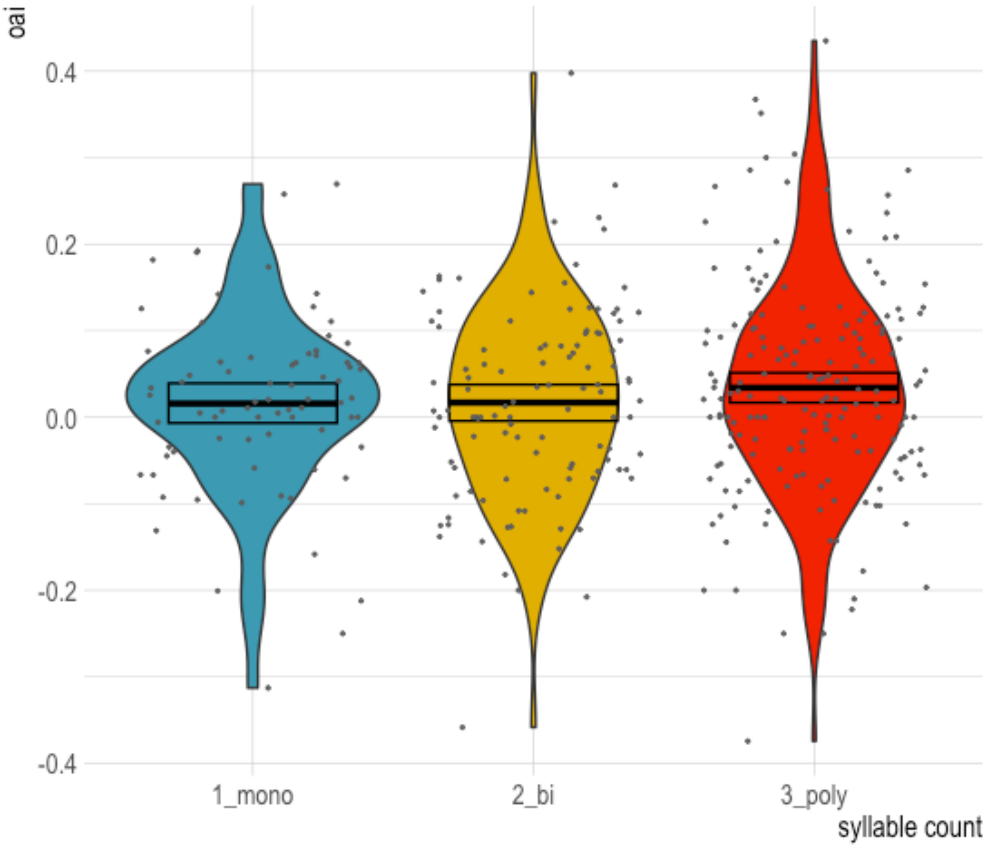


Figure 13 Violin plot showing distribution of OAIs for babble by utterance length in syllables and within-category variance. Bootstraps show category means and confidence intervals (0.95).

Models 6-7: modelling interactions between utterance complexity and age

Having found no significant effect of babble utterance length when laterality data from all points in developmental time were pooled, two further models were designed to test for interactions between age and utterance complexity as a measure of advancing articulatory skill. Again, it was not possible to construct single model testing for both one-to-one interactions since this would have rendered the resulting subsets of the babble OAI corpus small and vulnerable to unstable estimates.

Model 6: modelling age and babble utterance length

The first of these models was constructed as follows (5.18) and was tested against an interactional model excluding the random slopes (5.19) and a model excluding the fixed interaction (5.20) to identify the better fitting model:

(5.18) *lmer (oai ~ age_days * syllable_count + (1 + age_days * syllable_count | baby_name))*

(5.19) *lmer (oai ~ age_days * syllable_count + (1 | baby_name))*

(5.20) *lmer (oai ~ (1 + age_days * syllable_count | baby_name))*

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.18)		*-305.66*	*-196.16*
(5.16)	1	**-544.48**	**-513.19**
(5.17)	1	-350.60	-260.65

Table 25 Results of ANOVA comparing full age_days * syllable_count interactional model to a model excluding the random slopes (age_days * syllable_count) (5.16) and a model excluding the fixed age_days * syllable_count interaction (5.17).

The results of this exclusion process, shown in Table 25 , indicate the best fitting model to be one retaining the fixed interaction but excluding the random slopes (AIC = -544.48, BIC = -513.19) i.e., model (5.19). However, neither the fixed interaction nor the random slopes explain a significant proportion of the variance observed in these models (*Pr(>Chisq)* = 1). The results of this model are summarised in Table 26. The reference level (Intercept) is *bisyllabic babble.*

When age is at 0, the OAI estimate for bisyllabic babble utterances is marginally significantly different from 0 at -0.131 (*t* = -1.750, *df* = 0.0283, *p* = 0.081), showing right laterality, but this effect does not pass the p < 0.05 threshold for significance. For every one-unit increase in *age_days* (i.e., each passing day) the OAI for bisyllabic babble increases significantly by +0.0004784 indicating a significant leftward shift (*t* = 2.004, *df* = 0.0295, *p* = 0.46, see Appendix III: Model 5 for full output).

When age is at 0, the OAI estimate for monosyllabic babble is lower than the estimate for bisyllabic babble at -0.224, indicating stronger right laterality, but this difference is not significant (t = -0.730, *df* = 0.0362 p = 0.466). The estimate for monosyllabic babble was predicted to increase by +0.0002951 each day, indicating a leftward shift in laterality (*t* = 0.729, *df* = 0.0362, *p* = 0.4663), though this shift is not significant.

When age is at 0, the OAI estimate for polysyllabic babble is higher than the estimate for bisyllabic babble at -0.05, though, again, this difference is not significant (t = 0.806, *df* = 0.0359, *p* = 0.421). The OAI for polysyllabic babble is estimated to decrease by -0.0001934 per day, indicating a nonsignificant rightward shift (*t* = -0.613, *df* = 0.0357 *p* = 0.540).

As in stated with reference to Model 4, adults have been found to show stronger left lateralisation when producing longer sequences of oral gestures. The lack of a between-length difference here does not align with this finding. The estimated and actual trajectories in lateralisation over time of OAIs for utterances of each length (visualised in Figure 14) may offer some interesting insight, however. Notably, actual OAIs for monosyllabic babble were more strongly left lateralised than actual OAIs for either bisyllabic babble or polysyllabic babble, and actual OAIs for polysyllabic babble were more robustly left lateralised than those for either monosyllabic or bisyllabic babble. This observation will be explored further in Chapter 6. Further analysis of the effects of utterance length with a larger sample of still-frames or babies may be required confirm or rule out the presence of a differential effect of utterance length on the laterality of babble.

The marginal R^2 indicates that 3.5% of the variance observed in this model can be attributed to the fixed interaction between *age* and *syllable_count*, and the conditional R^2 indicates that the fixed and random predictors combined account for 4.4%.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
bi (Intercept)†	-0.131	0.075	-0.28 – 0.2	-1.750	**0.081**
mono	-0.093	0.127	-0.34 – 0.16	-0.730	0.466
poly	0.079	0.098	-0.11 – 0.27	0.806	0.421
age_days * bi	0.001	0.000	0.00 – 0.00	2.004	**0.046**
age_days * mono	0.000	0.001	-0.00 – 0.00	0.729	0.466
age_days * poly	-0.000	0.000	-0.00 – 0.00	-0.613	0.540
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby_name	0.00				
ICC	0.01				
^{N} baby_name	8				
Observations	369				
Marginal R^2	0.035				
Conditional R^2	0.044				

Table 26 Results of mixed effects model testing for differential effect of age on the laterality of monosyllabic, bisyllabic, and polysyllabic babble utterances. Reference level is bisyllabic babble. Significant p-values are shown in bold.

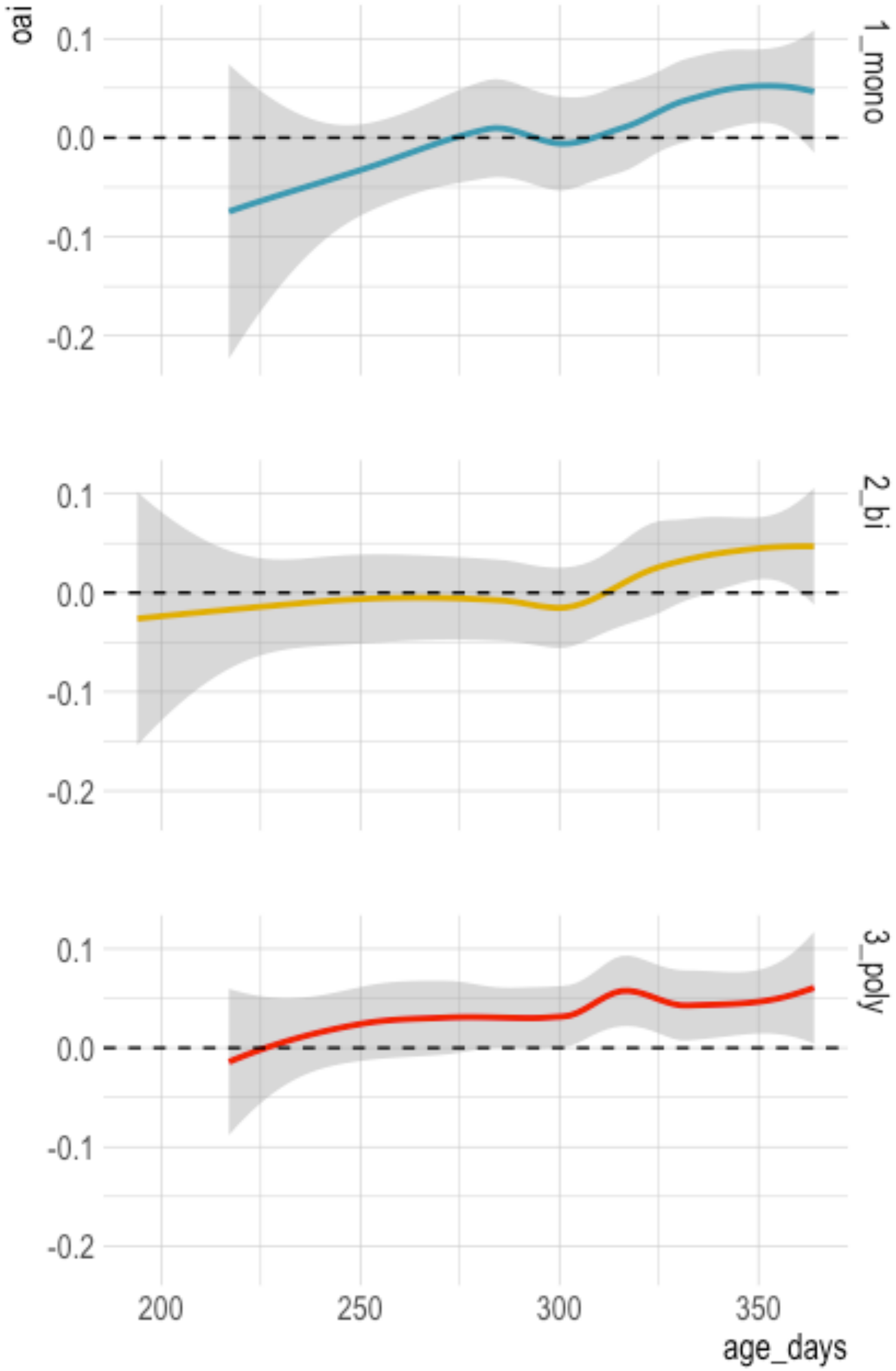


Figure 14 Spaghetti plot illustrating actual trajectories of OAIs for monosyllabic, bisyllabic, and polysyllabic babbles over time. Shaded areas indicate real-time variability. Positive values indicate left laterality and vice versa.

Model 7: modelling age and babble subtype

A final model was constructed to test for an interactional effect of *age* and *babble_subtype.* In his stage model, discussed in Chapter 1, Oller (1980) describes how singleton, reduplicated, and variegated babble represent increasingly sophisticated behaviours: singleton babble requires creating one close-open transition, reduplication involves repeating a close-open gesture, and variegation requires the baby to form two or more closures at different locations in the vocal tract. Oller states that some, though not all, babies progress through these behaviours gradually and in sequence. With reference to the laterality of hand dominance, Ramsay (1980, 1984) has suggested that some process of lateralisation may be associated with the emergence of reduplicated and variegated babble. Movement of the mouth and hands show evidence of activity in overlapping networks in the brain (Häberling & Corballis, 2016; Häberling et al., 2016). If the emergence of reduplicated and/or vareigated babble *is* associated with some neural reorganisation as well as the evident reogranisation in vocal behaviour, then it is possible that this may be reflected in the laterality of the orofacial gestures that produce these subtypes of babble. It is possible that the change in the laterality of babble as a function of developmental time observed in Model 2 may be associated with the emergence of the more sophisticated articulatory planning and control required to produce reduplicated and vareigated babble. An initial model was constructed as follows (5.21) and was tested against an interactional model excluding the random slopes (5.22) and a model exlcuding the fixed interaction (5.23):

(5.21) *lmer (oai ~ age_days * babble_subtype + (1 + age_days * babble_subtype | baby_name))*

(5.22) *lmer (oai ~ age_days * babble_subtype + (1 | baby_name))*

(5.23) *lmer (oai ~ (1 + age_days * babble_subtype | baby_name))*

Model	*Pr (>Chisq)*	*AIC*	*BIC*
(5.18)		*-342.51*	*-233.01*
(5.19)	1	**-544.07**	**-512.78**
(5.20)	1	-507.27	-417.32

Table 27 Results of ANOVA comparing full age_days * syllable_count interactional model to a model excluding the random slopes (age_days * syllable_count) (5.19) and a model excluding the fixed age_days * syllable_count interaction (5.20).

The results of this exclusion process, shown in Table 27, indicate the best fitting model to be one retaining the fixed interaction but excluding the random slopes (AIC = -544.07, BIC = -512.78). However, neither the fixed interaction nor the random slopes explain a significant proportion of the variance observed in these models (*Pr(>Chisq)* = 1). The results of this model are summarised in Table 28. The reference level (Intercept) is *reduplicated babble.*

No significant interactions between *age* and *subtype* were observed. When all predictors are at 0, the OAI estimate for reduplicated babble utterances is -0.052 (*t* = -0.527, *df* = 0.0343, *p* = 0.598), showing right laterality, but this estimate is not significantly different from 0. Every one-unit increase in *age_days* (i.e., each passing day) results in a non-significant leftward shift of +0.0002365 to the estimated OAI for reduplicated babble (*t* = 0.732, *df* = 0.0347, *p* = 0.465), see Appendix III: Model 6 for full output).

When age is at 0, the OAI estimate for singleton babble is more right lateralised than the estimate for reduplicated babble at 0.214 (*t* = -1.330, *df* = 0.0363, *p* = 0.184), though this difference is not significant. With each day, the OAI for singleton babble was estimated to increase by +0.0005329 indicating a non-significant leftward shift in laterality (*t* = 1.351, *df* = 0.0363, *p* = 0.177).

When age is at 0, the estimate for variegated babbles is less right lateralised than the estimate for reduplicated babble at -0.026 (*t* = 0.244, *df* = 0.0362, *p* = 0.823). Again, this difference is not significant. The OAI for variegated babble was estimated to decrease by -0.00005826 per day, indicating a non-significant rightward shift (t = -0.151, df = 0.362, p = 0.88).

The trajectories of actual OAIs for subtype of babble over time are visualised in Figure 15. Notably, actual OAIs for variegated babble show largely robust left laterality from day c.250 (see Figure 14), similar to actual OAIs for polysyllabic babble (Figure 13). Actual OAIs for reduplicated babble show the most real-time variability, whereas singleton babbles were initially highly variable but became less so over developmental time. As in the case of Model 5, further study with a larger sample and perhaps a longer data collection period may be helpful to confirm whether longer and more complex utterances show a sustained and significant rightward trajectory as a function of developmental time.

The marginal R2 indicates that 3.3% of the variance observed in this model can be attributed to the fixed interaction between age and subtype, and the conditional R2 indicates that the fixed and random predictors combined account for 4.8%.

Predictors	*Estimates*	*Std Error*	*95% CI*	*t*	*p*
redup (Intercept)†	-0.052	0.099	-0.25 – 0.14	-0.527	0.598
sing	-0.162	0.122	-0.40 – 0.08	-1.330	0.184
var	0.026	0.118	-0.21 – 0.26	0.224	0.823
age_days * redup	0.000	0.000	-0.00 – 0.00	0.732	0.465
age_days * sing	0.001	0.000	-0.00 – 0.00	1.351	0.177
age_days * var	-0.000	0.000	-0.00 – 0.00	-0.151	0.880
Random Effects					
σ^2	0.01				
$^{\tau}$00 baby_name	0.00				
ICC	0.01				
^{N} baby_name	8				
Observations	369				
Marginal R^2	0.033				
Conditional R^2	0.048				

Table 28 Results of Results of mixed effects model testing for differential effect of age on the laterality of singleton, reduplicated, and variegated babble utterances. Reference level is reduplicated babble.

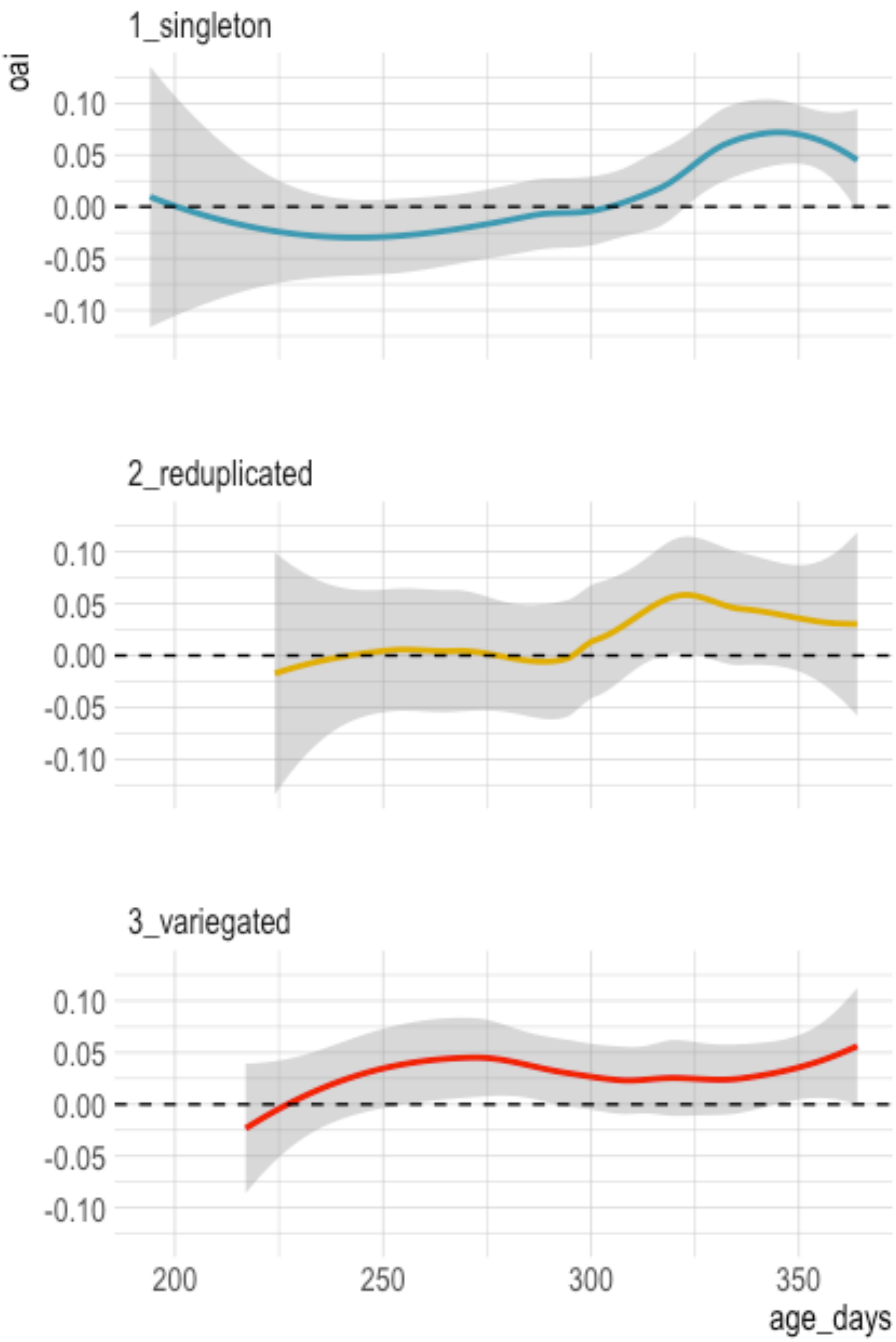


Figure 15 Spaghetti plot illustrating actual trajectories of OAIs for singleton, reduplicated, and variegated babbles over time. Shaded areas indicate real-time variability. Positive values indicate left laterality and vice versa.

General discussion

To summarise, Models 1-7 have reported the following findings. Model 1 indicated that babble is significantly left lateralised ($p < 0.001$) when OAIs for babble from all babies across the entire data collection period are pooled. Further to this non-babble and smiles and laugh show significantly different laterality to babble ($p < 0.001$). The predictors *age_days, vms,* and *sex* were excluded from this preliminary additive model because they did not explain a significant proportion on the variance observed in this model. Model 2 showed a significant interaction between babble and developmental time such that babble showed significant right hemisphere dominance at emergence ($p = 0.001$) and shifted leftward significantly over time, becoming left lateralised by the end of the data collection period ($p < 0.001$). Model 3 showed no significant interaction between babble and Vocal Motor Scheme attainment. Model 4 showed no significant sex difference in babies estimated OAIs for babble. Female babies showed significant left laterality ($p = 0.018$) and significant between category differences from non-babble ($p =0.017$) and smiles and laughs ($p = 0.000$). The estimated babble OAI for male babies was not significantly different from that for female babies ($p = 0.798$). Model 5 showed no significant between-length difference in the laterality of monosyllabic, bisyllabic, and polysyllabic babble. The predictor *subtype* was excluded from this additive model because it did not explain a significant proportion on the variance observed therein. Model 6 indicated a significant interaction between bisyllabic babble and developmental time, such that bisyllabic babble showed marginally significant right laterality at emergence ($p = 0.081$) and shifted leftward significantly as a function of developmental time, becoming left lateralised by the end of the data collection period ($p = 0.046$). Finally, Model 7 found no significant interaction between babble subtype and developmental time. The findings specifically pertaining to babble are summarised in Table 29 below.

That babble overall shows time-dependent change in both the direction and the strength of its lateralisation points to a role for post-natal articulatory experience in the lateralisation of babble. Since babble is a crucial prerequisite to speech and spoken language, this finding indicates that not all of the neural mechanisms involved in language are left lateralised from or before birth, as has been suggested by authors including, for example, Witelson and Pallie (1973), Chi et al. (1977), Wada et al. (1977), and Petitto and colleagues (1991, 2002a, 2004). The fact that bisyllabic babble becomes significantly left lateralised indicates a possible role for sequentiality (i.e., utterance length) in stimulating leftward lateralisation for babble. Possible reasons why OAIs for monosyllabic babble and polysyllabic babble did not show similar trajectories will be explored in Chapter 6.

The low Marginal and Conditional R^2 values generated by Model 1-7 indicate that these models contain a high proportion of unexplained variance. That is to say, the predictors in these models predict less variance than can be attributed to random variance. The random predictor *baby_name* was included in these models on the assumption made in psychological sciences that the data would contain more inter-individual variation than intra-individual variation (Gelman & Hill, 2006). However, this was not the case – all babies showed considerable variability with regard to the strength and direction of the asymmetry in their lip openings.

These findings have important implications and their discovery here demonstrates a clear advantage of both hemimouth area measurement and mixed effects modelling for exploring baby behaviour. Previous infant laterality research has given rise to the impression that 5-12-month-olds' babble may already be consistently and stably left lateralised and this assumption has been cited to support the Nativist perspective on phonological development (see Petitto and colleagues, 1991, 2002a, 2004). The findings of Models 1-7, possible explanations, and some theoretical implications will be discussed in Chapter 6.

Model	*Key findings*	*Significance*
Model 1	Babble showed significant left laterality	**$p < 0.001$**
	Significant between-category difference for babble and non-babble	**$p < 0.001$**

	Significant between category difference for babble smile/laugh	**$p < 0.001$**
Model 2	Babble showed significant right laterality at emergence	**$p = 0.010$**
	Significant interaction between babble and developmental time	**$p < 0.001$**
Model 3	No significant interaction between babble and Vocal Motor Scheme attainment	$p = 0.570$ (pre) $p = 0.740$ (at) $p = 0.078$ (post)
Model 4	Female babies showed significant left laterality for babble	**$p = 0.018$** (female)
	No significant interaction between babble and sex i.e., no significant difference in laterality	$p = 0.798$ (male)
	Bisyllabic babble showed non-significant left laterality	$p = 0.173$
Model 5	No significant between-length difference for monosyllabic and bisyllabic babble	$p = 0.972$
	No significant between-length difference for bisyllabic and polysyllabic babble	$p = 0.235$
Model 6	No significant interaction between monosyllabic babble and developmental time	$p = 0.466$
	Bisyllabic babble showed marginally significant right laterality at emergence	$p = 0.081$
	Significant interaction between bisyllabic babble and developmental time	$p = 0.046$
	No significant interaction between polysyllabic babble and developmental time	$p = 0.540$
Model 7	No significant interaction between babble subtype and developmental time	$p = 0.177$ (singleton) $p = 0.465$ (reduplicated) $p = 0.880$ (variegated)

Table 29 Summary of findings of Models 1-7 pertaining specifically to babble. Significant results are shown in bold.

Chapter 6 Discussion and Conclusions

"All the facts possible about the architecture of the end state won't tell you how it got there"

(Thelen & Smith, 1994, p. 36)

"…babbling is an exploratory stage in vocal development marked by many variable and repetitive vocalizations, for which production does not require a specific social or functional context, suggesting exploration…"

(Ter Haar et al., 2021, p. 2)

Introduction

The remainder of this thesis will consider how the findings presented in Chapter 5 speak to the questions posed in Chapter 4 and the literature reviewed in the chapters preceding. The implications of these findings for our understanding of babble and language ontogeny and phylogeny are also explored. This thesis has argued that Dynamic Systems theory offers a comprehensive, consistent, and convincing account of how babble emerges, how it changes with time and articulatory experience, and, eventually, how it furnishes the child with the tools to begin to do language. A new hypothesis of the underlying nature of babble – the Old Parts, New Machine hypothesis – has been put forward. It is argued in this chapter that this hypothesis offers a closer fitting explanation for the patterns of lateralisation observed in the present study than either the Linguistic hypothesis or the Motoric hypothesis. The contributions, affordances, and limitations of this study and the method used herein are discussed, and it is recommended that future orofacial asymmetry research should prefer this method over other less fine-grained or objective methods for analysing laterality and lateralisation via lip openings during orofacial gestures of interest. Throughout this chapter, directions for future research are also suggested.

Building on previous research

The first research question posed in this thesis concerned whether Holowka and Petitto's (2002a) finding of significant between-category differences in the laterality of babble, non-babble vocalisations, and smiles could be replicated using a more fine-grained approach for analysing orofacial asymmetry and using linear mixed effects modelling. Model 1 indicates that this is, indeed, the case: babble is significantly differently lateralised to non-babble and smiles or laughs. Model 1 gave an estimated OAI of 0.026 for babble, indicating significant left hemisphere dominance ($p < 0.001$). The OAI estimates for non-babble and smiles and laughs suggest weaker asymmetry and right hemisphere dominance at -0.003 and -0.009 respectively. Here, the results of the present study diverge from those of Holowka and Petitto (2002a).

Holowka and Petitto (2002a) found smiles and to be significantly right lateralised, whilst non-babble showed near equal hemispheric involvement, with a slight tendency towards left laterality. In the present study, neither non-babble nor smiles and laughs showed a significant degree of right laterality – that is, the OAI estimates generated by Model 1 for these categories of expression were close to 0. Considered alongside findings from Schuetze and Reid's (2005) baby and toddler laterality study, the findings of the present study suggest that emotional facial expressions may not stably lateralised in infancy and early childhood (see also Fox & Davidson, 1986; 1988; Lindell, 2017). Schuetze and Reid (2005), found 18-24-month-olds' laterality indices for smiles to be predominantly positive, indicating *left* laterality. However, the authors note that their mean laterality index for smiles across all babies sampled of 0.02 (*SD* = 0.07) did not show a significant degree of lateralisation when compared with their stronger mean laterality score of -0.15 (*SD* = 0.09) for facial expressions of distress (Schietze & Reid, 2005)[33].

Non-babble vocalisations were used in the present study and by Holowka and Petitto as a control category. It was expected that non-babble vocalisations would show high variability in laterality resulting in a mean category estimate close to zero, owing to this category containing quite a diverse range of orofacial gestures (squeals, growls, yawns, coughs, sneezes etc.) and this was confirmed to be the case. The more interesting finding is that confidence intervals and figures presented throughout Chapter 5 revealed similarly high quantities of variability in OAIs for babble (0.01 – 0.04), non-babble (-0.04 – -0.01), and smiles and laughs (-0.05 – -0.02) alike.

Babies in Holowka and Petitto's (2002a) study showed remarkable consistency in laterality, whilst the OAIs generated in the present study showed great variability both within and across babies and within and across home visits (see Figures 9 and 10). Table 30 below shows Holowka and Petitto's mean Laterality Indices (LIs). Each still-frame was coded as +1 (left lateralised), -1 (right lateralised), or 0 (equal hemispheric involvement), and mean LIs were generated for each baby by calculating the average of their LIs. The quota of still-frames per category per child is not stated, though given the stated total of 150 still-frames and assuming equal distribution, this may mean that five still-frames were analysed per category per baby. This being the case, a mean LI of +0.8 for babble would mean that four out of five still-frames showed left laterality, whilst one still-frame showed equal hemispheric involvement. If this is so, then 44/50 (88%) of Holowka and Petitto's (2002a) babble still-frames showed left laterality and 41/50 (82%) smile still-frames showed right laterality. In comparison, in the present study, 213/369 (57.72%) babble still-frames showed left laterality and 286/540 (52.96%) smile still-frames showed right laterality. There are several possible reasons for these differences relating to the methodological decisions undertaken.

---

[33] Schuetze and Reid's (2005) statistical analyses were carried out using ANOVAs, whereas mixed effects modelling was used in the present study. This may also partially explain why an estimated laterality index of 0.02 does not constitute a significant effect in Schuetze and Reid's study while an estimated OAI of 0.026 *does* constitute a significant effect in the present study.

Group	Baby	Gender	Age	Babble LI	Non-babble LI	Smile LI
English	E1	male	0;10.0	0.8	0	-1
English	E2	male	0;5.26	1	0	-1
English	E3	female	0;11.10	1	-0.2	-0.4
English	E4	female	0;12.0	1	-0.2	-0.8
English	E5	female	0;12.4	0.8	0	-1
			Mean LI	*0.92*	*-0.08*	*-0.84*
French	F1	female	0;9.27	0.8	-0.2	-0.6
French	F2	female	0;8.0	0.8	0	-1
French	F3	female	0;8.13	1	0	-1
French	F4	male	0;10.2	0.8	0	-0.6
French	F5	male	0;9.1	0.8	0.2	-0.8
			Mean LI	*0.84*	*-0.08*	*-0.80*
			Mean LI (all)	0.88	-0.08	-0.82

Table 30 Mean LI scores for each category of orofacial gesture for each baby in Holowka and Petitto's (2002a) study, reproduced from their supplementary material (Holowka & Petitto, 2002b) and modified for accessibility.

Firstly, in the present study, still-frames of babies laughing were included in the smile/laugh category. It is not stated whether the still-frames of smiles used by Holowka and Petitto (2002a) were silent smiles or whether any of these co-occurred with laughs. However, if activity in the right lateralised networks supporting emotional processing does significantly influence the strength of the asymmetry in the way that the mouth moves during facial expressions and speech, as it has been found to in several orofacial asymmetry studies (Graves et al., 1982; Wylie & Goodale, 1988; Schuetze & Reid, 2005; Nagy, 2012), then the inclusion of laughs – vocalised or aspirated expressions of emotion involving smiling – in this category should not result in significantly weaker or less frequently right lateralised OAIs.

Secondly, differences between Holowka and Petitto's (2002a) findings and the findings of the present study may arise from differences in the way that orofacial asymmetry was assessed. Holowka and Petitto (2002a, following Graves et al., 1982) used two rulers, placed parallel to the lips, to measure the angle created by the opening of the lips and used the direction of this angle to rate the direction of the asymmetry (+1, -1, 0). This method is highly cost-effective and allows for large amounts of data to be processed quite rapidly relative to some other methods. However, for its affordances, it also possesses certain limitations. Firstly, this method cannot tell us about the strength of any observed asymmetry. Secondly, even when a still-frame is greatly enlarged (Graves et al., 1982; Wyler et al., 1987) it may be difficult to identify subtle asymmetries by visual assessment alone. Thirdly, the lips are not straight lines. Some babies (and adults and children) have fairly linear lips, but others have more rounded mouths, and some have a marked peak or bow beneath the philtrum. Individual differences in orofacial physiology may mean that the rulers need to be placed in different locations relative to the lips for different still-frames and/or different individuals. Finally, the human mouth sometimes makes quite irregular shapes, and it may not always be clear where rulers should be placed. Since 5-12-month-olds are exploring their vocal apparatus and since their neuromuscular control over these is still developing, irregular mouth shapes may occur more frequently and may be more exaggerated amongst babies than mouth shapes seen in speech and facial expressions in adults. Rulers may be limited in their ability to accurately plot asymmetry in lip openings in examples like those in Figure 16 below. Hemimouth area measurement is much more time-consuming, but it allows the researcher to accurately trace the shapes made by the opening of the lips, revealing subtle differences in the distance between the lips at each point in their inner perimeter that may not be captured by rulers. Calculating OAIs using hemimouth area measurements returns information not only about the direction of asymmetry but also about the degree of asymmetry, gauging the strength of any asymmetry observed. The sensitivity of the method used in the present study may explain the higher within-individual and within-category variability and smaller effect sizes observed in this study than in Holowka and Petitto's study.

Figure 16 Top L: Benji shows prominent median tubercle (non-babble). Top M: Arthur shows rounded mouth shape (non-babble). Top R to bottom R: Adelaide (babble), Benji (non-babble), Cameron (babble), and Fred (non-babble), show irregular mouth shapes.

Thirdly, the longitudinal design of the present study introduces a further source of variation. Holowka and Petitto (2002a, 2002b) collected video data from babies only once, at the “exact” time when rhythmic, mature-sounding reduplicated babble first emerged (2002b, p. 1). In the present study, babies were visited repeatedly over the course 8 months, from a time before neuromuscular control and rhythmic ability are typically reliably established (Thelen, 1979, 1981; Ejiri, 1998; Ejiri & Masataka, 2001; Iverson et al., 2007), until the time when some babies begin to produce their first meaningful words (Oller, 1980; Stark, 1980; Studdert-Kennedy, 1990; Vihman, 2014). This is a time of substantial behavioural and neural reorganisation (Oller et al., 1976; Oller, 1980; McCune & Vihman, 1987; Van Der Knaap & Valk, 1990; Hayakawa et al., 1991; Thelen & Smith, 1994; Ejiri, 1998; Bogolepova & Malofeeva, 2001; Ejiri & Masataka, 2001; Keren-Portnoy et al., 2005; Paterson et al., 2006; Mareschal et al., 2007; Su et al., 2008; Tierney & Nelson, 2009; Perani et al., 2011; Petanjek et al., 2011; Vihman, 2014; Paredes et al., 2016; Williamson & Lyons, 2018; Oller et al., 2019). Some of this neural and behavioural reorganisation involves changes in the laterality of behaviours and cognitive processes like physical movement, socio-emotional, semantic and phonological processing, though some do not become stably lateralised until later in childhood (Ramsay, 1980, 1984; Best & Queen, 1989; Rothbart et al., 1989; Mundy et al., 2000; Schuetze & Reid, 2005; Workman et al., 2006; Dehaene-Lambertz et al., 2010; Nagy, 2012; Rosselli et al., 2014; Emerson et al., 2016; Hodgson et al., 2016; Lindell et al., 2017; Weiss et al., 2018; Reynolds et al., 2019; Yang et al., 2019; Olulade et al., 2020; Wang et al., 2021). The findings of the present study suggest that the neural network or networks supporting babble also undergo/es some process of hemispheric specialisation as a function of developmental time, which is reflected in the way that the jaw, lips and tongue are innervated. Whilst Model 1 estimates babble to be left lateralised when data from all ages is pooled together (0.026), Model 2 estimates babble to be right hemisphere dominant at emergence (-0.128), gradually shifting becoming increasingly left hemisphere dominant by 12 months of age.

Finally, whilst Holowka and Petitto (2002a, 2002b) focused exclusively on reduplicated babble, the present study examined the laterality of singleton, reduplicated, and variegated babble. In the present study, babble subtype was used to examine how many consonants were produced within an utterance. Babble utterances containing one single initial, medial, or final consonant were coded as *singleton* e.g., [da], [ib], or [ɛnɛ]. Babble utterances containing a single repeated initial or final consonant were coded as *reduplicated* e.g., [vev], [dɪdɪd], or [bɒbɒbɒbɒ]. Babble utterances containing more than one consonant with more than one place of articulation were coded as *variegated* e.g., [bɪð], [vʊdi], or [utigəʃɛw]. The exclusion process for Model 5 indicated that no significant proportion of variance in OAIs for babble could be attributed to babble subtype, suggesting that there was no significant between-subtype difference in laterality. Model 7 indicated that there was no significant change in laterality in any subtype of babble as a function of developmental time. It is therefore unlikely that including data from all three subtypes can explain this difference.

Instead, the findings relating to utterance length may offer more insight. As another measure of utterance complexity, babble utterances were coded as *mono, bi*, or *poly* according to the number of syllables that they contained. There was no significant between-length difference in laterality, with Model 5 yielding estimates of 0.017 for monosyllabic babble, 0.018 for bisyllabic babble, and 0.016 for variegated babble. However, the confidence interval for monosyllabic babble was wider (-0.03 – 0.03) and the standard error slightly higher (0.17) than those for bisyllabic (*CI* = -0.01 – 0.04 *SE* = 0.013) and polysyllabic babble (*CI* = -0.01 – 0.04, *SE* = 0.014), suggesting that monosyllabic babble was less stably lateralised than bisyllabic or polysyllabic babble. Additionally, a significant interaction with age was found in Model 6, such that that bisyllabic babble, initially showing marginally significant right laterality (*p* = 0.081), became significantly more left lateralised as a function of developmental time (*p* = 0.046). The estimated OAI for monosyllabic babble was not significantly difference from the estimate bisyllabic babble (*p* = 0.466) though its estimated OAI was more strongly negative than the estimate for bisyllabic babble. This suggests that monosyllabic babble also showed marginally significant right laterality at emergence. Unlike bisyllabic babble, however, there was no interaction between monosyllabic babble and developmental time. By contrast, the estimated OAI for polysyllabic babble was less strongly negative than that for bisyllabic babble. However, this estimate too showed no significant difference from the OAI estimates for bisyllabic (*p* = 0.421), suggesting that polysyllabic babble too showed marginally significant right laterality at emergence. Like monosyllabic babble, polysyllabic babble showed no interaction with developmental time. Holowka and Petitto (2002a, 2002b) do not comment on the number of syllables in the babble utterances from which still-frames were analysed in their study. Since they were investigating reduplicated babble, it is possible that their sample contained fewer or no monosyllabic utterances and higher proportions of bisyllabic and polysyllabic babble – both of which showed a stronger (though not absolute) tendency towards left laterality than did monosyllabic babble in the present study.

The laterality of babble

Like the laterality of speech, the laterality of babble can provide a window onto its underlying nature. The following section will consider what the findings of the present study may be able to tell us about how babble emerges and develops, and will consider three competing theories about the underlying nature of babble. One theory – the Linguistic hypothesis – posits that babble shares a neural substrate with language and is the manifestation of babies' innate drive to do language (see Petitto and colleagues, 1991, 2002a, 2004). Alternatively, Frame/Content theory – one conceptualisation of the Motoric hypothesis – suggests that babble shares a neural substrate with nutritive jaw oscillations and emerges when babies experiment with adding vocalisation to an already-stable motor production routine (see Davis and MacNeilage, 1993, 1995; MacNeilage, 1998).

An alternative theory – the Old Parts, New Machine hypothesis, based in Dynamic Systems theory – is put forward in this thesis. This hypothesis posits that babble emerges endogenously as the synergistic reorganisation of several diverse, asynchronously emerging, domain-general vocal and motoric capacities that are manifested in exploratory vocal and motor behaviours (like vocalising, mouthing, and rhythmic movement, amongst others) that come before it (e.g., see Oller, 1980; Iverson & Thelen, 1999; Ejiri & Masataka, 2001; and Chapter 1 of this thesis for a full discussion). This hypothesis further posits that babble becomes relevant to later language acquisition only once babies become proficient at producing consonant-vowel syllables. Once proficient, babies can begin to notice matches between the syllables that they themselves produce and the syllables that their caregivers and others around them produce. At this point, an innate, generalised, survival-based drive to behave like conspecifics and a domain-general capacity for social learning – evidence of both of which can be seen in early behaviours like vocal turn-taking (c.2-3 months) and babies' and children's imitative smiling (c.1-3 months) or clapping (c.9 months) (Vihman, 2014) – may drive babies to co-opt babble as another vehicle for social integration and bonding. This thesis argues that the Old Parts, New Machine hypothesis offers the best fitting conceptualisation of the underlying nature of babble both for the findings of the present study and for the findings of previous research as well. The following discussion will explore how the Old Parts, New Machine hypothesis aligns with the hypotheses and findings of the present study.

Right lateralisation and babble

Two key contributions of the present study are the findings that the laterality of babble shows inconsistency in real-time and that it changes significantly in both direction and strength as a function of developmental time. In another longitudinal laterality study, Schuetze and Reid (2005) found evidence of strengthening right laterality in facial expressions of distress in 12-24-month-olds and a pilot to the present study found strengthening left laterality in the babble of 9-18-month-olds (Oxley et al., 2014). To this author's knowledge, the present study is the first to find evidence that babble shows significant *right* laterality around the time of emergence and gradually shifts significantly leftward, becoming left hemisphere dominant, before 12 months of age (Model 2). Further to this, in the present study, hemimouth area measurement and mixed effects modelling have revealed considerable real-time variability in OAIs, indicating greater instability in the laterality of babble, which has not been observed before in any study to this author's knowledge (Models 1-7). So why might babble be right hemisphere dominant when it first emerges?

Between the beginning of the third trimester and around the sixth month post-natally, babies hear more vocalisations than they produce, meaning that their experiences of vocalisation are largely perceptual (Chi et al., 1977; Previc, 1991; Bogolepova & Malofeeva, 2001; Dehaene-Lambertz et al., 2002; May et al., 2011; Vihman, 2014). Further to this, babies' pre-natal perceptual experiences and early post-natal productive experiences of vocalisation mostly centre around modulation of pitch and voice quality (e.g., muffled sounds perceived during gestation, vocalic shouts, shrieks, growls) (Oller, 1980; Oller, 2000; Dehaene-Lambertz et al., 2002; May et al., 2011). Additionally, it can be observed in everyday caregiver-child interaction that many of the contingent and non-contingent vocalisations that young babies hear and produce carry emotional affect (e.g., infant directed speech and singing, laughs, cries, excited shrieks, caregiver commiseration (Meins et al., 1998; Gros-Louis et al., 2006, 2014; Vihman, 2014; Warlaumont et al., 2014). These types of auditory information are typically processed more efficiently by right lateralised neural networks for emotional processing in adults and older children (Johnson, 2000; Lindell, 2006; Mareschal et al., 2007; Morillon et al., 2010; Minagawa-Kawai et al., 2011; Sussman, 2015; Lindell, 2018).

Although the infant brain is organised differently to more mature brains, much of the neuronal migration that takes place during development does so during gestation (Thelen & Smith, 1994; Johnson, 2000; Mareschal et al., 2007; Sussman, 2015). Neurons that migrate to the right hemisphere during gestation tend to be better equipped at the neurochemical level to handle global information like pitch and emotion than neurons in the left (Johnson, 2000; Mareschal et al., 2007; Morillon et al., 2010; Sussman, 2015). The abundance of auditory-vocal information that can be processed efficiently by the right hemisphere during gestation and early infancy may explain why the right inferior frontal gyrus shows faster structural differentiation and cell proliferation during the first 12 months than its left homologue (Bogolepova & Malofeeva, 2001). The inferior frontal gyri and their homologues in non-human primates are implicated in interpreting complex auditory and visuospatial information and coordinating movements of the hands and mouth (Leiner et al., 1993; Fadiga et al., 1995; Iverson & Thelen, 1999; Vihman, 2002; Iverson & Fagan, 2004; Müller & Basho, 2004; Nishitani et al., 2005; Fadiga & Craighero, 2006; De Carli et al., 2007; Iverson et al., 2007; Skipper et al., 2007; Taglialatela et al., 2008; Petrides & Pandya, 2009; Häberling & Corballis, 2016; Häberling et al., 2016; Hodgson et al., 2021). Of particular relevance here, in adults, the right inferior frontal gyrus is activated as well as the left when we perceive and produce emotional vocalisations, and the relative strength of activation in each homologue varies as a function of whether an emotional vocalisation contains language with emotional prosody or is purely non-linguistic (Frühholz & Grandjean, 2013).

Repeated activation strengthens neural networks, resulting in preservation of synapses and, eventually, in myelination (Van Der Knaap & Valk, 1990; Mareschal et al., 2007; Petanjek et al., 2011; Williamson & Lyons, 2018). These things increase the likelihood that a neural network will be reactivated in future under similar conditions. Once a behaviour and its neural substrate become stable, it may be produced with less effort and modified more flexibly in real time without prompting any change in the overall form of the behaviour or the architecture of the substrate (Thelen & Smith, 1994; Mareschal et al., 2007; Vihman, 2014). A well-established, efficient right hemispheric neural substrate for vocalisation with a high likelihood of re-activation may act as a stable neural and behavioural attractor when babble first emerges (Thelen & Smith, 1994). By 6 months of age, a right lateralised network or networks for perceiving and producing vocal behaviour involving neurons in the right inferior frontal gyrus may be more well-established than any left hemispheric homologue that may also exist at this time (Bogolepova & Malofeeva, 2001). Such a substrate may become preferentially co-activated and connected with clusters of neurons supporting the capacities to plan and produce sequential and rhythmic motor behaviours, giving rise to syllabic CV alternations (Thelen, 1979, 1981, 1989; Ejiri, 1998; Bogolepova & Malofeeva, 2001; Ejiri & Masataka, 2001; Tierney & Nelson, 2009). This may be why babble was estimated and observed to show right hemisphere dominance at emergence in the present study (-0.0127, Model 2). Its laterality may be influenced and constrained by a pattern of cerebral organisation that is the result of a baby's dynamic history of experiences with vocalisation up until the time when the capacities to plan and instantiate rhythmic and sequential movements become stable and flexible enough to be combined with other in-repertoire behaviours.

Left lateralisation and babble

The significant age-related leftward shift in the laterality of babble observed in Model 2 may reflect an experience-dependent process of hemispheric specialisation. This thesis does not seek to argue that maturational growth plays no role in neural development – only that with the passage of time necessarily come repeated and novel experiences - we do not exist in a vaccum. These experiences necessarily form part of a baby's dynamic history, shaping and constraining their development in the future (Thelen & Smith, 1994). This experience-dependent specialisation may have the effect of optimising the efficiency of information transmission and the conservation of energy in the brain (Mareschal et al., 2007). Less energy is required to transmit electrical signals through shorter pathways between proximal clusters of neurons within one hemisphere than is required to transmit signals more distally or across hemispheres (Vallortigara et al., 1999; Mareschal et al., 2007; Corballis, 2008; Lindell, 2013; Sussman, 2015).

Rhythmic ability and sequential processing are specialised capacities that are more efficiently supported by neurons that typically migrate to the left hemisphere of the brain during gestation (Johnson, 2000; Mareschal et al., 2007; Morillon et al., 2010; Sussman, 2015; Hodgson et al., 2016; Hodgson et al., 2021). Being available from very early in development, it seems likely that phonation and the articulatory and cognitive flexibility required for voice modulation may be more generalised capacities, achievable with near-equal proficiency by either hemisphere upon which more complex behaviours like speech or singing can later build (e.g., see Tierney & Nelson, 2009; Minagawa-Kawai et al., 2011). In human and non-human animals alike, vocalisation can be recruited into networks that occupy either hemisphere depending on the emotional, social, or cognitive functions of other components of the networks in which it participates (Graves et al., 1982; Wolf &

Goodale, 1987; Wylie & Goodale, 1988; Graves & Landis, 1990; Hook-Costigan & Rogers, 1998; Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Siniscalchi et al., 2008; Minagawa-Kawai et al., 2011; Wallez et al., 2012; Siniscalchi et al., 2016; Siniscalchi et al., 2018).

When rhythmic and sequential ability become coupled with the capacity for vocalisation in babble production, less well-established clusters of neurons in the *left* inferior frontal gyrus may begin to be recruited to support the aforementioned right lateralised network/s with the work involved in modulating breath, vocal fold vibration, and supralaryngeal articulation serially rather than globally (Serrien et al., 2006; Sussman, 2015). When presented with novel challenges and functional limitations, existing neural networks are sometimes stimulated to recruit support from neuronal clusters better equipped to process information of the type being experienced (Serrien et al., 2006; Killgore & Yurgelun-Todd, 2007; Raja Beharelle et al., 2010; Olulade et al., 2020). The more frequently these supporting neuronal clusters are co-activated with the network/s supporting babble, the more likely it is that they will be re-co-activated in future (Thelen & Smith, 1994; Serrien et al., 2006; Mareschal et al., 2007). Eventually, the pre-existing network/s supporting babble may reorganise to incorporate these supporting left-hemispheric neuronal clusters in a more permanent way.

This reorganisation may give rise to a redundant network/s, with bilateral areas supporting the vocalisation component of babble. A body of research has provided evidence that language-related functions in infants and children show more bilateral representation and redundancy than adults, with selective pruning and reorganisation taking place gradually between infancy and adulthood (Thelen & Smith, 1994; Paterson et al., 2006; Perani et al., 2011; Petanjek et al., 2011; Rosselli et al., 2014; Emerson et al., 2016; Reynolds et al., 2019). Having two or more redundant and overlapping networks supporting babble may contribute to the rather high real-time variability in the laterality of babies in the present study (see Figure 9). The network selected to produce babble in any given moment may be influenced by dynamic processes in the brain, body, or environment, like spreading or residual activation from neighbouring brain regions, comfort level or affective state, interacting with toys or others (Fox & Davidson, 1986, 1988; Spivey, 2008). The adaptive reorganisation described here may also explain the long period of differentiation and cell proliferation observed in the left inferior frontal gyrus between 6-24 months (Bogolepova & Malofeeva, 2001). During this period, the left inferior frontal gyrus catches up with, and surpasses its right homologue in terms of cell density and relative size.

As babies consolidate their ability to produce syllables through repeated productive and perceptual experience, these syllables become more available, automatic, stable, and flexible (Thelen & Smith, 1994; Vihman, 2014). That is, syllables may be retrieved faster, may be produced with less effort, may have more reliable and consistent acoustic and articulatory results, and may be produced in different phonotactic or environmental contexts without prompting a reorganisation of their acoustic or articulatory form. It is possible that the strengthening and growth observed in the left inferior frontal gyrus by Bogolopova and Malaofeeva (2001) may be associated with increasingly frequent activation of neuronal clusters in the left inferior gyrus supporting the instantiation, storage, and retrieval of motor plans for vocalisations comprising regularised CV transition timing. These things together may explain why, in the present study, monosyllabic babble was estimated to shift leftward gradually, eventually becoming robustly left lateralised, as a function of developmental time, albeit non-significantly (Model 6, Figure 14). The lack of a significant effect for monosyllabic babble may stem from the fact that monosyllables, involving very short sequences of articulatory gestures compared with longer utterances, may draw less on typically left hemispheric aptitudes for interpreting, constructing, and instantiating sequential motor patterns (see Petrides & Pandya, 2009; Buchsbaum et al., 2001; Hickok & Poeppel, 2007; Mesgarani et al., 2014; Corballis, 2015)

Utterance complexity

Left lateralisation and babies' vocal exploration during babble likely interact cyclically, mutually influencing and constraining one another (Thelen & Smith, 1994). With advancing time and articulatory experience, babies begin to produce an increasingly diverse range of consonants and they develop the ability to repeat or combine these in new ways (Oller et al., 1976; Oller, 1980; McCune & Vihman, 1987; Oller, 2000; Vihman, 2014). Research with adults suggests that producing longer utterances or a greater variety of articulatory gestures may draw more heavily and more frequently on left hemispheric neuronal clusters that are more adept at doing sequential vocalisation than their right homologues. The incorporation into the babble network/s of these supporting left hemispheric clusters may furnish babies with a strengthened capacity to plan and instantiate longer and more complex sequences. This, in turn, may result in the more frequent production of longer utterances containing a wider range of more stable consonants that is typically observed from around 10 months of age (Oller et al., 1976; Oller, 1980; Ramsay, 1980; Stark, 1980; Ramsay, 1984; McCune & Vihman, 1987; Studdert-Kennedy, 1990; Stoel-Gammon, 1992; Oller, 2000; Vihman, 2014).

Once well-established, left lateralised networks may come to be preferentially activated during babble. Changes in myelination associated with repeated coactivation may mean that left hemisphere structures involved in babble receive stronger, faster signals from the left hemispheric neuronal clusters responsible for rhythm and serial processing (Van Der Knaap & Valk, 1990; Hayakawa et al., 1991; Vallortigara et al., 1999; Mareschal et al., 2007; Corballis, 2008; Su et al., 2008; Lindell, 2013; Sussman, 2015; Williamson & Lyons, 2018; Grotheer et al., 2022)). Over time, synapses connecting the leftward-shifting babble network/s to structures in the right hemisphere that support globally modulated vocalisation (e.g., pitch variation) may decay, reducing the likelihood of re-co-activation, and increasing the likelihood that these connections will eventually be pruned (Mareschal et al., 2007; Petanjek et al., 2011). Thus the strengthening of the leftward asymmetry may be magnified by decreasing right hemisphere involvement in babble. This may explain why the left inferior frontal gyrus shows signs of cell proliferation and synaptic strengthening between 6-24 months of age (Bogolepova & Malofeeva, 2001). The following two sections will consider how this explanation may account for the differences in trajectories for different subtypes of babble and babble utterances of different lengths.

Utterance length

No previous study, to this author's knowledge, has investigated the relationship between laterality and babble utterance length, so these findings represent a further contribution of the present study to the fields of laterality and pre-linguistic development.

Modelling between-length differences

Adult humans show stronger left laterality when producing longer and/or more complex sequential vocalisations, regardless of whether these sounds carry any meaning (Wolf & Goodale, 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). The vocal-motoric component of longer utterances requires more detailed and serialised articulatory planning and effort, and this means that neural architecture in the left hemisphere, adept at sequential planning and action (Sussman, 2015), consumes more processing power (i.e., more blood flow and stronger electrical signals) during longer utterances than during shorter utterances (Salmelin & Sams, 2002; Saarinen et al., 2005). Unlike adults in previous research, babies in the present study showed no significant between-length difference in the laterality of their monosyllabic, bisyllabic, and polysyllabic babble utterances (Model 5). The estimated OAI for polysyllabic babble (0.034) did show somewhat stronger left laterality than the estimates for monosyllabic (0.019) and bisyllabic (0.018) babble but these effects were not significant (*p* = 0.972 and *p* = 0.235 respectively).

The lack of significance here is perhaps unsurprising given babies' marked inconsistency in laterality in real-time observed in Model 1 and the significant effect of age on the laterality of babble observed in Model 2. With respect to Model 5, it was anticipated that monosyllabic babble would show the weakest left hemisphere dominance (or even perhaps the strongest tendency towards right hemisphere dominance), and that polysyllabic babble would show the strongest left hemisphere dominance, with bisyllabic babble falling somewhere in between. As babies produce babble utterances with more syllables, it was theorised that neuronal clusters in the left hemisphere – possibly in the left inferior frontal gyrus – may come to be relied upon significantly more heavily and more frequently. Monosyllabic babble, presenting a shorter sequence and a lighter articulatory load, may require less processing power or less specialised left hemispheric support (Serrien et al., 2006; Killgore & Yurgelun-Todd, 2007; Raja Beharelle et al., 2010; Vihman, 2014; Olulade et al., 2020). Polysyllabic babble involves more sequential processing and presents a weightier load, so may draw more heavily on left hemispheric resources. This prediction was not borne out in Model 5 and so, in this respect, the findings of the present study do not align with the findings of adult research concerning sequential vocalisation (Wyler et al., 1987, Salmelin & Sams, 2002; Saarinen et al., 2005). It should be noted that OAIs for babble utterances of all lenghts showed considerable variability to the extent that the range occupied by OAIs for utterances of each length overlapped with one another, as indicated by the confidence intervals for monosyllabic (*CI* = -0.03 – _0.03), bisyllabic (*CI* = -0.01 – _0.04), and polysyllabic (*CI* = -0.01 – 0.04) babble.

There are two possible explanations for this non-significance. The first is that the construction and instantiation of long sequences may be a function that becomes localised to the left hemisphere later on in the lifespan between infancy and adulthood (see Workman et al., 2006; Lindell et al., 2017; Weiss et al., 2018; Wang et al., 2021 for other behaviours showing lateralisation persisting into childhood). The second is that the present study employed too small or unbalanced a sample of still-frames extracted from babbles of each length to observe any statistically meaningful between-length difference (monosyllabic babble n = 78; bisyllabic babble n = 112; polysyllabic babble n = 156; nTotal = 369). To this end, future research with a larger sample of participants or still-frames and/or a longer data collection period may help to identify at what point in the lifespan length of utterance begins to affect the laterality of an utterance significantly and reliably, as it does in adults.

Modelling interactions between utterance length and developmental time

Model 6 indicates that there was a significant interaction between utterance length and age for bisyllabic babble. The estimated OAI of -0.131 for bisyllabic babble when it first emerged is negative but does not indicate significant right laterality, (*p* = 0.081). Bisyllabic babble was estimated to shift leftward significantly over developmental time (*p* = 0.046) becoming left lateralised before 12 months of age (Figure 14 illustrates the trajectory of actual OAIs for all lengths of babble).

No such interaction was seen for monosyllabic or polysyllabic babble. Monosyllabic babble was more strongly right lateralised than bisyllabic babble when it first emerged with an estimated OAI of -0.224, though the between-length difference was not significant (*p* = 0.466). While Model 6 predicts OAIs for monosyllabic babble to increase over developmental time indicating a slight leftward shift (see Model 6) this effect does not pass the threshold for significance (*p* = 0.466). Polysyllabic babble was more weakly right lateralised at the time of first emergence, with an estimated OAI of -0.05. Again, Model 6 indicates that this estimate was not significantly different from the estimate for bisyllabic babble (*p* = 0.421). OAIs for polysyllabic babble were estimated to decrease indicating rightward shift over developmental time, though Figure 14 indicates that all actual OAIs for polysyllabic babble were > 0, indicating robust left laterality from c.225 days of age. The apparently regressive rightward shift predicted by Model 6 for polysyllabic babble may reflect higher within-category variability in OAIs for polysyllabic babble between emergence and c.240 days of age and the plateau-peak-decline in OAIs between c.275-330 days depicted in Figure 14 rather than an actual overall rightward shift.

It has been suggested that many (though not all) babies begin with shorter, simpler babble utterances and that longer and more complex sequences emerge only once they can produce single syllables reliably (Oller, 1980; Stark, 1980; Elbers, 1982; Elbers & Ton, 1985). It was anticipated, therefore that, at the point when polysyllabic babble emerged, some process of left lateralisation might already be underway as a result of babies drawing on more left hemispheric processing resources (e.g., see Sussman, 2015) when producing bisyllabic babble. It was expected that this might mean that polysyllabic babble might show less exaggerated change in laterality over developmental time compared with monosyllabic and bisyllabic babble. The gradual leftward trajectories of actual OAIs shown in Figure 14 seem superficially to support this idea. However, it should be noted that the analyses and visuals reported here contain only data relating to babble utterances for which still-frames were available and suitable for analysis. This was determined by whether the baby was facing the camera and remained sufficiently still with the mouth unobstructed and clearly visible during a vocalisation. As such, this information may not represent the actual distribution and order of emergence of babble utterances of different lengths for babies in the present study.

The significant leftward trajectory in the estimated OAI for bisyllabic babble observed in Model 6 ($p = 0.046$) also partially aligns with this expectation. What was unexpected was that neither monosyllabic babble nor polysyllabic babble showed any significant leftward shift. One possible explanation for this is that monosyllabic babble may require relatively little specialised left hemispheric processing and so does not need to draw on increasing left hemisphere involvement. Instead, a more bilaterally distributed may suffice to carry out the work of constructing and instantiating single syllables. Indeed, such a network may suffice for simple, one-movement oral gestures throughout the lifespan. This would align with findings that, in adults, producing meaningless isolated syllables is associated with relatively weak left laterality compared with producing longer sequences (Wyler et al., 1987; Salmelin and Sams, 2002; Saarinen et al., 2005).

Another possibility is that the laterality of babble follows a pattern of non-linear change during infancy, becoming stably and robustly left lateralised only later in life. The very low Marginal R^2 (0.035) and Conditional R^2 (0.044) for Model 6 indicate marked inconsistency within OAIs for each length of utterance and within each baby. Non-linear and sometimes apparently regressive behavioural changes are characteristic of development in biological organisms, including humans (Thelen & Smith, 1994). During infancy, the brain contains multitudinous connections capable of performing the same or very similar functions and which network is activated in any given moment may be partially determined by what other processes are taking place in the brain concurrently (Fox & Davdison, 1986; Mareschal, 2007; Petanjek, 2011). The adult-like pattern of increased left laterality for longer or more complex vocalisations and mouth movements may emerge later.

To explore these possibilities, the present study could be extended by investigating what proportion of each baby's babble at each home visit was monosyllabic, bisyllabic, or polysyllabic. The laterality of babble utterances could be traced over a longer timescale, and the polysyllabic babble category, which here included any babble of more than three syllables in length, could be further divided to allow for more fine-grained longitudinal comparisons of the laterality of utterances of different lengths. The effects of utterance length could also be explored in babies and toddlers early monosyllabic, bisyllabic, and polysyllabic words.

Acquiring and articulating consonants

Previously, some laterality research has found evidence to suggest that acquiring consonants and combining them in new ways may be associated with hemispheric specialisation in babies (Ramsay, 1980, 1984; Oxley et al., 2014). In the present study, babble subtype (singleton, reduplicated, or variegated) did not explain a significant proportion of the variance observed and so this predictor was excluded from Model 5. No significant interaction was observed between subtype and age (Model 7). The estimates for singleton (*OAI* = -0.215, *p* = 0.184), reduplicated (*OAI* = -0.052, *p* = 0.598), and polysyllabic (*OAI* = -0.026 *p* = 0.823) were negative at the time of babble emergence, but these estimates did not indicate significant right lateralisation .

OAIs were estimated for singleton (*p* = 0.177) and reduplicated babble (*p* = 0.465) were estimated to increase as a function of developmental time, whilst the OAI for variegated babble decreased (*p* = 0.880) but these effects did not show significant change over time. As with polysyllabic babble though, Figure 15 shows that actual OAIs for variegated babble were fairly consistently above 0 indicating left laterality, and more so than actual OAIs for singleton and reduplicated babble. The apparent rightward regression predicted by Model 7 may reflect the peak-decline-plateau in left laterality observed between c.275-320 days (Figure 15). Actual and estimates OAIs for babble subtype show a similar non-significant pattern of lateralisation to utterance length, with more complex utterances showing a weaker tendency towards right laterality at emergence and a stronger tendency towards left laterality towards the end of the first year of life than less complex utterances. However, as noted, in this case, no significant interaction with age was observed.

As stated earlier, it has been theorised that babies first produce only singleton babble, with reduplicated babble emerging later, and variegated babble later still (Oller, 1980; Ramsay, 1980; Stark, 1980; Elbers, 1982; Ramsay, 1984; Elbers & Ton, 1985; Oller & Eilers, 1988; Studdert-Kennedy, 1990; Ejiri, 1998; Oller, 2000; Ejiri & Masataka, 2001). It was therefore anticipated that singleton and reduplicated babble might show more exaggerated leftward shift than variegated babble, because it was expected that variegated babble might be more left lateralised at the point of emergence than reduplicated or singleton babble. This expectation was not borne out in the data analysed here, since no significant interaction between age and any subtype of babble was observed.

In the present study, singleton babble was available for analysis from c.195 days of age, before either of the other subtypes. However, the next emerging subtype appears to have been variegated babble at c.220 days, whilst the first reduplicated babble in the sample is dated to c.225 days. In the light of the prediction made here, this might explain why there is no significant difference in the trajectories of OAIs over time for reduplicated and variegated babble. What this would not explain is why there is no significant difference between the trajectory seen for OAIs for singleton babble and either of the other subtypes of babble. In any case, this idea should be viewed with the caveat that these graphs include only data relating to the still-frames analysed and so do not represent the actual distribution of babble utterances of each subtype in the corpus as a whole. As with utterance length, the present study could be extended by investigating this idea further.

Alternatively, it may be that the neural reorganisation that takes place when babies acquire new consonants does so in a part of the brain that does not directly affect the way that the mouth moves. Speech sounds are represented in neuronal clusters in the superior temporal gyrus (Buchsbaum et al., 2001; Chang et al., 2010; Mesgarani et al., 2014). The superior temporal gyrus is not, however responsible for constructing and instantiating the motor routines involved in producing rhythmic sequential vocalisations. Instead, this is handled by the inferior frontal gyrus, to which the superior temporal gyrus is connected (Hickok & Poeppel, 2007). This may explain why we see an interaction between bisyllabic babble and developmental time – as babies become more and more proficient at producing mature-sounding syllabic vocalisations with time and practice (Vihman, 2014) – but not for babble subtype and developmental time.

Again, issues relating to sample size and distribution may contribute to the lack of a significant effect observed here. The babble corpus of nTotal = 369 still-frames was divided n = 133 still-frames of singleton babble, n= 80 of reduplicated babble, and n = 156 of variegated babble. Alternatively, as suggested with reference to utterance length, the left hemispheric propensity for sequential vocalisations and mouth movements may not have stably lateralised by the age of 12 months and the differentiation in left hemisphere dominance for simpler vs. more complex mouth movements and vocalisations may emerge later in childhood or adolescence (e.g., see)

Utterance length and babble subtype - implications

Because utterance length and babble subtype provide measures of different kinds of utterance complexity, another key contribution of the present study is the finding that the length of a sequential utterance but not the consonants it comprises can effect significant changes in the laterality of babble. To this author's knowledge, no other orofacial asymmetry study has sought to discriminate between the way in which number of syllables vs. number of consonants may affect laterality either in adults or in children. As a result, these findings may offer some insight to an ongoing debate concerning the minimal unit of language processing. A body of research has sought to establish whether the syllable or the phoneme should be classified as the minimal unit of language (e.g., Morillon et al., 2010; Mahmoudzadeh et al., 2013). Simultaneous fMRI (functional Magnetic Resonance Imaging) and EEG (Electroecephalography) research with adults has identified neural activity associated with syllabic but not phonemic parsing in those exposed to natural language (Morillon et al., 2010), whilst babies show neural activity associated with both simultaneously when exposed to isolated words (Mahmoudzadeh et al., 2013). These findings may be associated with the quantity and structure of the auditory information perceived. Alternatively, they may reflect differences in degree of cerebral specialisation and in what is salient to the individual as a result of these neural organisational differences and of past and present auditory and articulatory experiences. The findings of the present study may suggest that, for production, the capacity to combine syllables requires more left hemispheric work, lateralises earlier, or becomes more strongly lateralised than the capacity to repeat or combine supraglottal closures within an utterance. However, the lack of significant between-category differences in the laterality of monosyllabic, bisyllabic, and polysyllabic babble mean that this interpretation should be taken with caution. Again, extending the present study with a larger sample of participating babies could be helpful to confirm whether any such effect may exist in the wider population or to identify when babble or speech becomes stable enough for a significant effect of utterance length to emerge. Another interesting question is whether marginal babble and other vocal behaviours that pre-date babble, lacking regular rhythm and sequential structure, also show any evidence of laterality and lateralisation.

Vocal Motor Schemes

The present study’s findings relating to Vocal Motor Scheme attainment may offer further insight here. Research literature reviewed in Chapter 3 has found evidence to suggest that acquiring consonants may be associated with experience-dependent proliferation and differentiation in neuronal clusters in the superior temporal gyrus responsible for representing individual speech sounds (Buchsbaum et al., 2001; Chang et al., 2010;

Mesgarani et al., 2014). Unexpectedly, in the present study, attaining a Vocal Motor Scheme was not associated with any truly significant proportion of the variance observed in the laterality of babble – or of any other category – suggesting that developing a stable and reliable consonant production routine did not significantly effect a change in the overall laterality of all babble utterances. This is contrary to the findings of the pilot to this study (Oxley et al., 2014). This difference in findings may reflect that fact that babies in the pilot study were fewer in number (n = 2) and considerably older (9-18 months) than babies in the present study. Alternatively, this difference may arise from differences in the statistical methods used. The pilot study used t-tests whereas the present study used mixed effects models, which are more rigorous and can differentiate between random variance and meaningful variance. Another possibility is that the absolute number of different consonants in a baby's productive repertoire at any given point in time, stable or not, may be sufficient to induce left lateralisation. Another thing that neither the pilot nor the present study tested was whether babies showed any difference in laterality when producing utterances containing a Vocal Motor Scheme consonant compared with utterances containing only non-Vocal Motor Scheme consonants. New data could be added to data from the present study could be to extend the findings reported here by investigating these possibilities. At present, insufficient still-frames of babies producing babble utterances containing their Vocal Motor Schemes are available for analysis.

Alternatively, the localisation of speech sound representations to the left hemisphere may be the work of months and years of vocal experience, and speech sounds may initially be more bilaterally represented since babble, at emergence, showed right hemisphere dominance (Rosselli et al., 2014; Emerson et al., 2016; Reynolds et al., 2019). Until around 11.5 months, both the inferior frontal gyri and superior temporal gyri show asynchronous shifts towards functional *symmetry* rather than asymmetry (Emerson et al., 2016) associated with degeneracy and redundancy in neural organisation (Thelen & Smith, 1994). After this age, more pronounced asymmetries begin to emerge. The superior temporal gyrus and inferior frontal gyrus in the left hemisphere show asymmetrically increasing connectivity between 2-5 years of age (Reynolds et al., 2019), perhaps reflecting the gradual strengthening of a neural pathway for mapping acoustic signals onto articulatory routines (Buchsbaum et al., 2001; Hickok & Poeppel, 2007; Corballis, 2015). In light of this, the findings of the present study may suggest that, for production, the capacity to combine syllables into bisyllabic utterances may involve a greater degree of neural reorganisation, requiring more left hemispheric work than the move from producing utterances containing one consonant to utterance containing multiple different consonants, since bisyllabic babble shows significant change in laterality over time, whilst no significant lateralisation is observed for any subtype of babble (singleton, reduplicated, or variegated).

Lateralisation for babble and sex

Some research with adults and non-humans has found equivocal evidence of sex differences in lateralisation for language and vocalisation. In word retrieval and isolated word production, men and women have shown similar tendencies towards left lateralisation, but women have sometimes shown weaker left lateralisation than men in fluent speech or when experiencing amusement (Graves et al., 1982; Hausmann et al., 1998). However, when producing meaningless spoken syllables and non-speech mouth movements, women have shown stronger left lateralisation than men (Wolf & Goodale, 1987). As discussed in Chapter 3, this could reflect genuine differences in neural organisation in men and women. However, no significant structural or functional between-sex differences have been found by neuroimaging research to day (Frost et al., 1999; Clements et al., 2006; Wallentin, 2018).

An alternative possibility is that hormonal fluctuations may cause the brain to select different pathways at different times. The present study offered the opportunity to explore whether between-sex hormonal difference affect the laterality of babble in babies under 1 year of age.

During the exclusion process for Model 1, the predictor *sex* did not explain a significant proportion of the variance observed in this model and so was excluded. This implies that there was no significant difference between male and female babies' OAIs for babble, non-babble, or smiles and laughs in the present study. Likewise, Model 4 showed no significant difference in male and female babies' laterality for babble (p = 0.0.798).

A likely explanation for these findings is that hormonal differences between male and female babies are much smaller than hormonal differences between male and female adults or even adolescents. The hormone typically suggested to be associated with in fluctuations in laterality – oestradial – is the hormone associated with puberty and menstruation (Perret, 1986; Fernández et al., 2003; Lindell, 2006). Pre-menopausal women's blood contains 30-400 picograms of oestradial per millilitre. By contrast pre-pubertal girls' blood contains 1.6-2.6 picograms per millilitre and pre-pubertal boys' blood contains 0.4-0.11 picograms per millilitre. It may be interesting for future laterality research to investigate sex differences in laterality around the time of the onset of puberty when the concentration of oestradiol in the female body changes. It should be noted though, that as the participant sample involved in this study contained five boys and only three girls, the corpus of data available is unbalanced across sexes. The present study could be extended by the addition of data from more female babies to address this imbalance.

Laterality and the underlying nature of babble

The finding that babble shows evidence of left laterality has previously been interpreted as evidence to support the Linguistic hypothesis (Petitto and colleagues, 1991, 2022a, 2004). This hypothesis posits that babble is a "fundamentally linguistic" behaviour and the output of a "brain-based language capacity", triggered by exposure to the ambient language (Petitto & Marentette, 1991, p. 1945; Holowka & Petitto, 2002a; Petitto et al., 2004, p. 46). Under this hypothesis, babble is believed to emerge as babies' goal-directed attempts to produce language. The Linguistic hypothesis was constructed with the aim of addressing shortcomings of the Motoric hypothesis – to account for the 'speechiness' of babble, its apparently sudden emergence, and the widely observed continuity between babble and later language acquisition (Menn, 1971; Waterson, 1971; Oller et al., 1976; Priestly, 1977; Elbers, 1982; Elbers & Ton, 1985; Vihman et al., 1985; McCune & Vihman, 1987; Locke & Pearson, 1990; Locke, 2000; Keren-Portnoy et al., 2005; DePaolis et al., 2011; McGillion et al., 2017). According to the Linguistic hypothesis, since language is left lateralised, and babble – which is acoustically similar to spoken language and which furnishes the baby with phonological resources that are later useful to language acquisition – is also left lateralised, this indicates that babble and language share a universal domain-specific neural substrate (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004).

Left laterality and language

Earlier orofacial asymmetry research has found evidence of robust left laterality in the lip openings of adults during spoken word generation, syllable production, and fluent speech (Graves et al., 1982; Wyler et al., 1987; Graves & Landis, 1990). Further to this, purely linguistic speech tasks elicit more strongly left lateralised responses than speech tasks involving emotional or visuospatial processing (Graves et al., 1982; Wyler et al., 1987). Speech utterances involving more cognitive or articulatory planning or effort also show stronger left laterality (Wolf & Goodale, 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). These findings have been validated by studies using a range of behavioural techniques (Van Der Haegen & Brysbaert, 2018) and neuroimaging and neurosurgical methods (Deppe et al., 2000; Knecht et al., 2000; Buchsbaum et al., 2001; Hickok & Poeppel, 2007; Chang et al., 2010; Morillon et al., 2010; (Mesgarani et al., 2014); Nenert et al., 2017;). fTCD has found evidence of left lateralised blood flow associated with speech processing and linguistic cognition and fMRI has identified specific areas and networks primarily occupying the left hemisphere through which this blood flows when we interpret meaning and produce words (Deppe et al., 2000; Knecht et al., 2000; Buchsbaum et al., 2001; Chang et al., 2010; Nenert et al., 2017). A simultaneous fMRI and EEG paradigm has identified left lateralised expression of frequencies (‘brainwaves’) associated with interpreting syllables when we are exposed to natural language (Morillon et al., 2010). iEEG has identified clusters of neurons responsible for speech sound perception (Mesgarani et al., 2014) and projections have been found connecting these clusters to other clusters that are responsible for instantiating motor programmes associated with these sounds (Buchsbaum et al., 2001; Hickok & Poeppel, 2007; Corballis, 2015). Whilst some structures in the right hemisphere of the brain are involved in language-related processing, (Harpaz et al., 2009; Binder, 2017), these are fewer in number amongst most adults. So then, the evidence that language is supported by networks predominantly involving left hemispheric structures is convincing.

The present study has confirmed that Holowka and Petitto’s (2002) finding relating to the laterality of babble can be replicated using a more fine-grained method for analysing orofacial asymmetry (Model 1). This finding is not inconsistent with the Linguistic hypothesis. However, the Linguistic hypothesis is not the sole possible explanation as to why babble may be left lateralised and findings arising from laterality research in humans and non-human animals, as well as research from the fields of developmental psychology and neuroscience, do not align as comfortably with this perspective. Indeed, the present study has made two further key findings that *are* inconsistent with the Linguistic hypothesis. Firstly, the finding that babble is right hemisphere dominant at the time when it emerges and shifts leftward significantly as a function of developmental time (Model 2) raises a problem for the idea that babble is triggered and governed by a biologically pre-programmed, left hemispheric language capacity. Secondly, the finding that bisyllabic babble shows a significant interaction with developmental time (Model 5) suggests a role for experience-dependent neural reogranisation (see Thelen & Smith, 1994; Mareschal et al., 2007) and partially aligns with findings from adult research that complex sequential motor movements of the mouth (and hands) require more left hemispheric work than simple single articulatory gestures (Wyler et al., 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). For these reasons, this thesis argues against the Linguistic hypothesis and argues instead for the Old Parts, New Machine hypothesis.

Triggering by input

The suggestion that the onset of babble is triggered directly by perceiving the ambient language (Petitto and colleagues, 1991, 2002a, 2004) marginalises the critical role of directionless vocal exploration in babble emergence. Assuming triggering by language input has two possible interpretations. On the one hand, as suggested by Petitto and colleagues, a brain-based capacity for doing language may unfold at around 6-8 months by some as yet unidentified maturational process and, once in existence, may be activated when it receives linguistic input (Petitto & Marentette, 1991; Petitto et al., 2004). This language capacity, quite independently of any preceding productive or perceptual experience of vocalisation, may drive the baby to attempt to copy language, resulting directly in well-formed mature-sounding reduplicated syllables (Petitto & Marentette, 1991; Petitto et al., 2004). This interpretation is problematic. By the time that this language capacity is suggested to unfold, typically developing babies already have a number of varyingly sophisticated or well-formed, transitional, endogenously motivated vocal behaviours in repertoire (Oller et al., 1976; Oller, 1980; Stark, 1980; Oller, 2000; Vihman, 2014). As discussed in Chapters 1-2, these pre-babble exploratory behaviours refine and reinforce neuromuscular coordination and articulatory skills that babies appear to make extensive use of in babble, like creating and maintaining vowel resonance, forming supraglottal closures, and performing rhythmic motor actions. Like babble, babies may do these pre-babble behaviours reciprocally during interactions with caregivers from around 2-3 months, but, like babble, babies may also perform them spontaneously, independently of any interaction or, indeed, of any attempt to seek a caregiver's attention (Oller, 1980; Vihman et al., 1985; Oller, 2000; Vihman, 2014; Oller et al., 2019).

In the present study, six tokens of marginal babble were captured for Arthur and two for Leif. These utterances contained consonants but differed from babble in several ways. Consonants were not consistently produced in alternation with vowels, articulation was imprecise, reflecting weaker neuromuscular control, and the timing of the transition between the closed and open phases of CV articulations was not governed by regular, adult-like rhythm. Transcriptions of these marginal babble utterances are shown in Table 31 below. Upon consultation, Arthur and Leif's caregivers reported that they produced vocalisations of this type very rarely. Given the characteristic scarcity of marginal babble (Oller, 1980; Oller, 2000) it is quite possible that other participating babies also produced some such sounds but that these were not captured during the 30-minute twice-monthly recording sessions. The fact that babies already produce vocalisations like isolated consonants, vocalic shouts, or indeed marginal babble before the innate language capacity is proposed to mature raises the question of whether recourse to an innate language capacity is completely necessary in order to explain babble onset. Instead, babies might arrive at syllable production simply by combining behaviours already in repertoire along with a capacities for rhythmic and sequential movement in a more exploratory, haphazard way (Oller, 1980).

Baby	*Age (days)*	*Transcription*
Arthur	186	[gɪl]
Arthur	186	[βːəβ β β]
Arthur	186	[mːβwə]
Arthur	186	[hːə h h]
Arthur	186	[əːːɣi]
Arthur	186	[ədð dðdð]
Leif	181	[ʊvːː]
Leif	181	[iːːɛɱːːː]

Table 31 Marginal babble utterances produced by Arthur and Leif during each baby's second home visit

Another possible interpretation is that an innate left hemispheric neural substrate for language unfolds and is activated some time before 6-8 months of age and that this is what drives babies to gradually modify their vocal behaviour towards the adult model over the course of several months. Since, for typically hearing babies, spoken language input is available from birth and information about the prosody and rhythms of language is available prenatally, this triggering could potentially happen any time between birth and babble emergence. However, this idea would assume a degree of goal-directed behaviour for which we have little supporting evidence in 0-6-month-olds. Whilst babies may be observed to modify some of their vocal and motor behaviours from around 3-5 months (e.g., reaching, imitating facial expressions, vocal turn-taking), these modifications have short-term goals and involve only repetition or minor changes in the form of the behaviour (e.g., direction of reach, rounding/spreading of the lips) (Thelen & Smith, 1994; Vihman, 2014). By contrast, the journey between birth and babble emergence is marked by a much more substantial and indirect process of behavioural reorganisation involving many heterogenous and asynchronous behaviour modifications. Not until around 3-4 years of age, do children reliably demonstrate the ability to modify their behaviour to attain indirect goals (Klossek et al., 2008; Klossek et al., 2011; Taffoni et al., 2014). Even then, experience gathered through goal-free exploration plays a key role in their ability and inclination to do so (Taffoni et al., 2014).

Neither does triggering by input account for why deaf and hearing impaired babies and toddlers are observed to produce speechlike syllables in cases where spoken language input and auditory-vocal feedback are compromised or entirely inaccessible, albeit less frequently than their hearing counterparts (Oller & Eilers, 1988; Petitto & Marentette, 1991; Oller et al., 2019). Sensitivity to the ambient language input may seem to offer some explanation as to why, at the group-level, vowel and consonant production in babble may show some ambient language influence (de Boysson-Bardies et al., 1981; de Boysson-Bardies et al., 1989; de Boysson-Bardies & Vihman, 1991; de Boysson-Bardies, 1993). However, it does not account for cross-linguistic commonalities in the phonotactics of babble, nor for the high variability observed in the consonants favoured by individual babies in their babble and first words (i.e., their Vocal Motor Schemes) in spite of strong similarities in consonant distribution in caregiver speech amongst families with the same ambient language (de Boysson-Bardies et al., 1981; Locke, 1983; Elbers & Ton, 1985; Davis & MacNeilage, 1995; MacNeilage, 1998; MacNeilage & Davis, 2000; DePaolis et al., 2011; DePaolis et al., 2013; Majorano et al., 2014).

Davis and MacNeilage (1993, 1995, 1998) have put forward an alternative theory – Frame/Content theory – based in the Motoric hypothesis, which offers a more psychologically plausible account of how babble may emerge, grounded in the principle of descent with minor modification (MacNeilage & Davis, 1993; Davis & Macneilage, 1994; Davis & MacNeilage, 1995; MacNeilage, 1998). It is suggested that the jaw motion involved in babbling and speech (the *frame*) originates in nutritive jaw motion and that, over time, increasingly differentiated lip and tongue movements (the *content*) are added to these jaw movements. However, this theory is problematised by evidence that speech and nutritive jaw oscillation show different rates and trajectories of development (Wilson et al., 2008; Nip et al., 2009; Morrill et al., 2012) and have different oscillatory frequencies (Jürgens, 1998; Ghazanfar & Takahashi, 2014; Zimmerman & Foran, 2017). Further to this, evidence from behavioural and neuroimaging research suggests that speech and non-speech mouth movements are most likely supported by partially overlapping but distinct task-dependent neural networks (Jessen et al., 1999; Bunton, 2008; Watson & Lof, 2008; Wilson et al., 2008; Nip et al., 2009; Kent, 2015; Häberling & Corballis, 2016; Häberling et al., 2016; Kern et al., 2019; Parra-López et al., 2022). So, neither the Linguistic hypothesis nor the Motoric hypothesis can offer an entirely psychologically plausible account of how babble emerges.

Well-formedness

The present study has identified two further complications for both the Motoric and Linguistic hypotheses: namely that the laterality of babble is not stable in real time or developmental time. The estimated OAI for babble of 0.026 in Model 1 indicates left laterality, but the CI of 0.01-0.04 indicates real time variability (see Figure 9) and the marginal R^2 for this model indicates that category explains only a very small proportion of the variance observed here (1.4%). In Model 2, the interaction between babble and age is significant at $p = 0.001$. When age is at 0 (i.e., 5 months, being the age at which data collection began), the OAI estimate of -0.128 for babble indicates right hemisphere dominance, and babble is predicted to shift leftward significantly by +0.0004154 with each passing day.

Frame/Content theory posits that babble "emerges more or less fully formed rather than being put together in the uncoordinated manner characteristic of many baby action sequences (e.g., using culinary utensils)" (MacNeilage & Davis, 2001, p. 79). The findings of Models 1 and 2 suggest otherwise. The real-time variability observed here suggests that multiple degenerate and redundant networks may underlie babble (Thelen & Smith, 1994; Mareschal et al., 2007). That is, one function may be produced by a range of different structures, and each structure in the developing brain may be involved in a range of different behaviours depending on context and what neighbouring or connected structures are also concurrently activated. The architecture that is selected to support any given babble utterance may be determined by the speech sounds being targeted combined with dynamic contextual factors. Spreading or residual activation in the brain associated with emotion, cognition, sensory perception may increase the likelihood of one network or cluster of neurons receiving activation over another (Graves et al., 1982; Fox & Davidson, 1986; Wyler et al., 1987; Thelen & Smith, 1994). The initial right laterality of babble and its left lateralisation over developmental time runs counter to the claim that babble shows well-practiced or adult-like articulation from the time of its first emergence. These observations, coupled with babies' aforementioned steady transition through phasic pre-requisite behaviours to babble, provide further support for the claim that babble is in fact a dynamic, adaptive assembly of heterogenous vocal, rhythmic, and motor capacities rather than the simple re-purposing of an entrenched capacity for cyclical nutritive jaw movement (MacNeilage & Davis, 1993; Davis & MacNeilage, 1995; MacNeilage, 1998).

These findings also raise complications for the Linguistic hypothesis, which rests on the claim that left laterality is an indicator of linguistic-ness (Petitto & Marentette, 1991; Holowka & Petitto, 2002a; Petitto et al., 2004). One possible interpretation is that the neural substrate of babble *is* underlyingly linguistic but that this substrate is initially right lateralised rather than left lateralised. In this case, left laterality may not constitute strong or straightforward evidence that babble is fundamentally linguistic. Another possible interpretation is that babble is *not* supported by the same neural networks as language a) because these have not yet developed and b) because babble does not make use of all of the structures involved in language processing (e.g., neuronal clusters involved in syntactic, semantic, and pragmatic processing). Rather, babble may be supported by one or likely more, task-dependent network/s of domain-general clusters of neurons, which is/are lateralised differently to networks supporting language processing in adults. Because language and babble share articulatory and acoustic properties like phonation and supraglottal closure (Oller, 1980; Stark, 1980; Elbers, 1982; Elbers & Ton, 1985), the networks that support them may occupy overlapping but distinct territories in the brain (see also Häberling & Corballis, 2016; Häberling et al., 2016; Nenert et al., 2017). Whilst both of these interpretations are broadly in line with the fact that neural organisation in babies is not adult-like and is highly dispersed, redundant, and degenerate, the second interpretation is favoured here because it aligns more closely with current knowledge about the structural and functional organisation of the brain during development and during adulthood (Thelen & Smith, 1994; Serrien et al., 2006; Mareschal et al., 2007; Petanjek et al., 2011; Rosselli et al., 2014; Häberling & Corballis, 2016; Häberling et al., 2016).

Meaning, meaninglessness, and sequentiality

A particular problem for the Linguistic hypothesis is that non-linguistic orofacial movements are also left lateralised (Losin et al., 2008; Wallez et al., 2012). Chimpanzees, who do not possess language, show left hemisphere dominance when producing non-species-typical learned vocalisations in which gestures of the jaw, lips, and tongue must be combined with phonation in novel ways (Losin et al., 2008; Wallez et al., 2012). Marmosets – more distantly related to humans – show left hemisphere dominance when producing long, complex, neutral/affiliative vocal sequences (*twitter*) than when producing fear vocalisations in which only pitch is modulated (*tsik*) (Hook-Costigan & Rogers, 1998). In these cases, meaning is encoded in the rhythmicity and sequential coordination of articulatory gestures and the serial ordering of syllables or syllable-like segments in the absence of any innate capacity for language. Other primate species like Campbell's monkeys and putty nosed monkeys produce context-specific concatenated strings of vocalisations (Arnold & Zuberbühler, 2006; Ouattara et al., 2009; Zuberbühler, 2020). It may be of interest for future research to examine these monkeys' laterality in longer and shorter concatenated utterances.

Human adults show left hemisphere dominance when producing meaningless syllables (e.g., /bu/ and /bi/), silent vowel gestures (e.g., rounding or retracting the lips), and silent non—speech orofacial gestures (e.g., touching the teeth with the tongue) as well as in meaningful language (Wolf & Goodale, 1987; Wyler et al., 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). Further to this, stronger asymmetries are observed for repeated or serially-organised meaningless orofacial gestures than for isolated ones (Wolf & Goodale, 1987; Salmelin & Sams, 2002; Saarinen et al., 2005). Together, these findings suggest that the work involved in producing long and complex sequential orofacial gestures is sufficient to generate left lateralised activity in the absence of any linguistic processing. The fact that these studies generated complementary results using a range of methods (MEG, laterality rating, and lip displacement) and stimuli gives validation to this interpretation of their findings.

In the present study, bisyllabic babble utterances were estimated to show a weaker tendency towards right laterality at emergence, a stronger tendency towards left laterality at 12 months, and a lesser degree of within-category variability than monosyllabic babbles (Model 5, Figure 13). Bisyllabic babble was estimated to become significantly left lateralised over time (Model 5). Actual OAIs for monosyllabic, bisyllabic, and polysyllabic babble indicated left laterality during most of the data collection period (Figure 13). Babble, by definition, involves no encoding or interpretation of meaning into or out of sequential vocalisations nor the serial manner in which they are produced and so may have more in common with non-linguistic vocalisations and orofacial gestures like these than with language or vocal communication in non-human animals. What makes babble and other non-linguistic orofacial movements left lateralised, then, may be not some innate neural substrate for language, but rather the lateral distribution of a dynamic, synergistic, task-dependent neural network/s, composed of primarily left hemispheric neuronal clusters supporting lower level capacities like vocalisation, rhythm, and sequential processing. Producing more complex vocal sequences may involve greater left hemispheric activation in certain parts of this network than producing simpler or non-sequential vocalisations.

In light of this suggestion, it is also possible that the stronger asymmetries observed during complex and/or purely linguistic vs. emotional or visuospatial speech tasks (Graves et al., 1982; Wolf & Goodale, 1987; Wyler et al., 1987; Graves & Landis, 1990; Salmelin & Sams, 2002; Saarinen et al., 2005) may reflect not simply the increased activation in a brain based language capacity but rather may at least partially reflect proportionally greater reliance on typically left hemispheric capacities for constructing and instantiating long sequences of complex, rhythmic vocalisation.

Task-dependence and domain-generality

A fourth problem for the Linguistic hypothesis is that many other vocal and manual behaviours besides babble and language are connected with left lateralised neural activity and leftward asymmetries in cerebral size and sulcal and gyral morphology (e.g., Binder et al., 1996; Boradfield et al., 2001; Marie et al. 2018). Musical training and expertise are associated with leftward asymmetries in the superior temporal gyrus in humans (Glanville et al., 1977; Binder et al., 1996; Gannon et al., 1998; Keenan et al., 2001; Griffiths & Warren, 2002; Koelsch et al., 2005; Ono et al., 2011; Meyer et al., 2012; Elmer et al., 2016; Binder, 2017). Chimpanzees and baboons show leftward asymmetries in the size and density of the superior temporal gyrus, associated with the complexity of their vocal and social behaviours (Gannon et al., 1998; Poremba et al., 2003; Marie et al., 2018; Becker et al., 2021; Becker et al., 2022) and this leftward asymmetry has been found to increase as a function of developmental time in infant baboons (Becker et al., 2021). Endocasts and skeletal analysis have revealed evidence of leftward asymmetries in the size, organisation, and differentiation of the inferior frontal gyrus in several hominin ancestors who long pre-date the emergence of the capacity for modern human-like speech and language (Holloway, 1983; Tobias, 1987; Broadfield et al., 2001; Delson et al., 2001; Wu et al., 2006; Wu et al., 2011a; Wu & Pan, 2011b; Holloway et al., 2018; Dusseldorp & Lombard, 2021). Some authors have suggested that these asymmetries may have been associated with tool use and handedness as well as vocal communication.

Leftward asymmetry in the inferior frontal gyrus in chimpanzees, bonobos, orangutans, and gorillas as well as humans is implicated in interpreting, representing, retrieving, and coordinating complex manual movements and non-speech orofacial gestures in (Cantalupo & Hopkins, 2001; Vihman, 2002; Hopkins & Cantalupo, 2003; Crockford et al., 2004; Hecht & Parr, 2015; Häberling & Corballis, 2016; Häberling et al., 2016; Hodgson et al., 2016; Neubauer et al., 2020; Hodgson et al., 2021). Of particular significance are the findings that non-human primates show left hemisphere dominance when producing learned communicative vocalisations (Hook-Costigan & Rogers, 1998; Losin et al., 2008; Wallez et al., 2012).

Musical cognition, complex manual behaviours, and the vocal behaviour of non-human animals may seem to bear little superficial resemblance to babble. However, by considering commonalities between highly specialised left lateralised behaviours, we can learn about what lower level behaviours and capacities are foundational to them (Thelen & Smith, 1994; Tierney & Nelson, 2009). Having done so, we may then attempt to map these foundational behaviours and capacities onto the current understanding of how the brain works. From there, we can begin to formulate theories about the underlying nature of highly specialised behaviours.

The Old Parts, New Machine hypothesis

A theory of babble should attempt to account for all of the findings discussed here: how babble emerges and develops, how it influences the transition into language acquisition, and what left laterality may mean for our understanding of babble (and perhaps also of language). Findings from the present study, along with findings from laterality research with adults and comparative research with non-human primates, offer insights that are of value in attempting to formulate such a theory. These insights align more closely with the Old Parts, New Machine hypothesis than with the Linguistic or Motoric hypotheses. The Old Parts, New Machine hypothesis is grounded in the idea that new behaviours emerge in a non-teleological way as a result of phenomena in the distant or recent dynamic history of the organism (Thelen & Smith, 1994; Tierney & Nelson, 2009). It is argued that babble emerges as the synergistic product of a confluence of behaviours and capacities that pre-date it, some of which may have very ancient roots. It is further argued that babble becomes relevant to language acquisition only at such a time as babies produce utterances with sufficiently precise articulatory and rhythmic control to begin to notice matches between their own utterances and those of adults and older children around them. Babies are argued to arrive at this phase through endogenously motivated vocal exploration. Having arrived in this phase, the capacity for babble may be co-opted by an innate, domain-general drive to behave like conspecifics – an ancient evolutionary adaptation to ensure survival in social species (Donald, 1991). How this may manifest in the brain has been explored here. By working to understand how the brain connects domain-general clusters of neurons into highly specialised task-dependent networks in an adaptive, non-teleological way, this hypothesis aims to be sufficiently detailed to account for phenomena in developmental based on our current understanding of the brain, but also sufficiently flexible to be updated and modified as and when our understanding of the workings of the brain and body progresses.

Babble and vocal learning

Research concerning babble and vocal learning in non-human animals offers some interesting insights that are consistent with the claims made by the Old Parts, New Machine hypothesis but run counter to the Linguistic and Motoric hypotheses. Vocal learning refers to animals' gradual modification of the frequency, rhythm, serial ordering, and context of production of their vocal behaviour towards adult- or other-modelled targets (Janik &

Slater, 1997; Tyack, 2008; Carouso-Peck et al., 2021; Janik & Knörnschild, 2021).[34] Vocal learning is well-documented amongst humans, in babies' and adults' acquisition of words and grammar (Vihman, 2014). A growing body of research has found that some non-human animals, so far including seals, sea lions, whales, dolphins, elephants, bats, mice, mole-rats, goats, and some primates, also modify their vocal behaviour (Janik & Slater, 1997; Wilbrecht & Nottebohm, 2003; Tyack, 2008; Carouso-Peck et al., 2021; Janik & Knörnschild, 2021). The capacity for vocal learning other species serves various adaptive functions. Auditory signals can be used as 'names' to distinguish individuals or 'passwords' to signal group membership (Carouso-Peck et al., 2021; Janik & Knörnschild, 2021). In the infants of social species, the endogenous drive to vocalise represents a survival advantage. Vocalising when alone can elicit caregiver attention or can signal that the infant is healthy, safe from harm, and developing typically (Wallez & Vauclair, 2012; Oller et al., 2019; Carouso-Peck et al., 2021; Ter Haar et al., 2021). Intrinsic volubility also promotes the evolution and passing down of complex systems for communicative vocalisation used for social bonding, locating food, mating, and evading danger (Omedes, 1985; Elowson & Snowdon, 1994; Hook-Costigan & Rogers, 1998; Arnold & Zuberbühler, 2006; Knörnschild et al., 2006; Ouattara et al., 2009; Knörnschild, 2014; Fernandez et al., 2021; Ter Haar et al., 2021). That humans and other animals alike exhibit a capacity for vocal learning suggests that it may be quite an ancient trait, inherited from a common mammalian ancestor. Related research suggests that the drive to engage in vocal turn-taking similarly pre-dates modern humans, since it is shared by lemurs, marmosets, squirrel monkeys, Campbell's monkeys, and gibbons alike (Levinson, 2016).

Much vocal learning literature concerns birds (Pepperberg, 1998; Scanlan & Rogers, 1998; Ter Haar et al., 2021). Vocalisation in birds and mammals involves very different physiology and may represent a case of convergent evolution (Balter, 2010). In birds, the larynx is involved only in ingestion (c.f. Scanlan & Rogers, 1998). Instead, birds vocalise by exhaling through membranes in the *syrinx* (Snowdon, 2017). This organ is located at the tracheobronchial junction and is specifically evolved for phonation. It contains no flexible

---

[34] Janik and colleagues favour a narrower definition of vocal learning, excluding contextual learning. Following Tyack (2008), an expanded definition is favoured by this author because this allows us to consider a wider and richer range of commonalities between human and non-human vocal learning, allowing us to consider in more detail what common capacities underlie it.

tissue comparable to the vocal folds (Snowdon, 2017). Rather, sound is modulated by innervating muscles in the sides of the neck or chest simultaneously or independently, resulting in one or two simultaneous vocalisations. Unlike humans, bird vocalisations do not comprise any separable supralaryngeal modulation beyond the opening and closing of the beak (MacNeilage, 1998). Because of these anatomical differences and for brevity, the literature on birds is not discussed further here.

Some non-human mammals show quite flexible vocal learning throughout the lifespan like humans do. For example, captive seals and killer whales can create original vocalisation types by modifying in-repertoire vocalisations and may mimic their human caregivers' vocalisations (Carouso-Peck et al., 2021; Janik & Knörnschild, 2021). Three seals have been taught to 'sing' vowel sounds to the tune of *Twinkle, twinkle little star* (Stansbury & Janik, 2019). One Asian elephant has learned to imitate Korean vowel formants in order to bond with his caregivers by inserting his trunk into his mouth and vocalising nasally (Stoeger et al., 2012). Wild grey seal pups spontaneously modify the pitch contours of their earliest vocalisations to assimilate with other pups around them in the birthing colony, and they show faster learning when the targeted vocalisations are more consistent (Stansbury & Janik, 2021). These findings demonstrate a clear capacity for innovating vocalisations or modifying vocalisations towards a target sound – key facets of vocal learning.

The issue of vocal learning in non-human primates is somewhat more controversial: in the past, primate vocalisations were believed to be quite inflexible, innate, and involuntary. Some more research may provide evidence to the contrary (see Schel et al., 2013b). For example, pygmy marmosets in all stages of life develop flexible social group 'dialects' and show mutual convergence in the pitch and duration of *trill* vocalisations during prolonged exposure to new social groups (Elowson & Snowdon, 1994). Pygmy marmosets also voluntarily adapt their vocal behaviour in real time: individuals produce context-specific alert calls when they find a food source, but can choose to remain silent to avoid jeopardising a hunt when the food source that they find is live prey (Snowdon, 2018). Adult male pygmy marmosets selectively produce flexibly concatenated vocalisations comprising food calls and locational calls to signal their availability and location to prospective mates before but not after mating (Snowdon, 2018). Campbells' monkeys, who produce context-specific concatenated calls composed of as many as 25 discrete and varying vocalisations, show evidence of social vocal learning (Ouattara et al., 2009; Lemasson et al., 2011). Socially bonded individuals exhibit more similar articulation and concatenation patterns than genetically related individuals (Lemasson et al., 2011)

Wild chimpanzees vocalise intentionally and in a goal-directed way, targeting specific individuals (Schel et al., 2013a; Schel et al., 2013b) and conveying referential meanings in elicit specific reactions from others (Slocombe & Zuberbühler, 2005). This implies an understanding of cause and effect and an ability to adapt their behaviour to influence others. Adult and infant chimpanzees learn to modulate the volume, pitch, and duration in their vocalisations voluntarily to convey information about their affective state and social rank, and those of other chimpanzees around them (Slocombe & Zuberbühler, 2007). Like pygmy marmosets' *trills*, wild and captive chimpanzees' *pant-hoots* show between-group articulatory and acoustic divergences and within-group vocal convergence that can be attributed neither to genetic relatedness nor to similarities in living environment like tree density or availability of food (Mitani et al., 1992; Arcadi, 1996; Marshall et al., 1999). Neighbouring social groups who occupy overlapping territories show greater between-group divergence in the pitch and rate of their *pant-hoots* whilst social groups who live more distally show lesser divergence (Mitani et al., 1992). Some chimpanzees show vocal innovation, combining in-repertoire vocalisations to form new vocalisation types, the uptake of which throughout the group may depend on the social status of the innovator (Arcadi, 1996; Marshall et al., 1999). Captive chimpanzees have been trained to use attention getting sounds for communicating with their human caregivers (Losin et al., 2008; Taglialatela et al., 2008; Taglialatela et al., 2011; Wallez et al., 2012; Russell et al., 2013). Interestingly, the left inferior frontal gyrus in chimpanzees shows greater activation during heterospecific interaction in chimpanzees who have been taught attention getting sounds than amongst chimpanzees who have not (Taglialatela et al., 2021). These things point to a critical role for vocal innovation and cultural transmission in chimpanzees' acquisition and production of species typical vocalisations. These findings also demonstrate that the capacity for encoding and interpreting referential meaning in vocalisations is not the exclusive to humans.

Babble in non-human animals

The young of certain species, including bats, marmosets and perhaps also bonobos and whales, engage in vocal behaviours that bear behavioural, acoustic, and articulatory similarities to babble and so some authors have extended the term to these behaviours (Omedes, 1985; Elowson et al., 1998; Knörnschild et al., 2006; Knörnschild, 2014; Snowdon, 2018; Fernandez et al., 2021; Ter Haar et al., 2021). Ter Haar et al. (2021, p. 2) propose the following alternative conceptualisation of babble, which may be applied consistently across humans and certain species:

> "Babbling is an exploratory stage in vocal development marked by many variable and repetitive vocalisations for which production does not require a specific social or functional context, suggesting exploration […] Babbling is a precursor to the adult form of vocal communication in the sense that the sounds produced in babbling incorporate acoustic features required in the adult vocal system. In the most advanced babbling forms across various taxa, syllable-like elements emerge that often constitute exemplars of syllable-like elements of the adult system"

This conceptualisation allows us to recognise parallels between human and non-human babble that may inform how we think about the underlying nature of human babble, like its meaninglessness, repetition, exploratory-ness and (proto)syllabicity. That human and non-human babble exhibit common properties suggests that at least some of the neural architecture supporting babble may be shared across species. This further weakens the idea that an innate language capacity is necessary to explain babble emergence (see Petitto and colleagues, 1991, 2002a, 2004). Some of the foundational behavioural components of babble may long pre-date modern humans, let alone language. Instead, this conceptualisation lends further support to the central argument of the Old Parts, New Machine hypothesis – that human babble emerges as a synergy of pre-existing behaviours (e.g., the drives to vocalise and socialise), capacities (e.g., for voluntarily modifying vocalisations), and pieces of neural architecture (e.g., the inferior frontal gyrus and superior temporal gyrus) – by suggesting that certain components of babble predate it both in developmental and evolutionary time.

Babble in bats

From birth, greater sac-winged bat and greater spear-nosed bat pups produce isolation calls – individual-specific requests for caregiver attention, whose acoustic structure is determined by vocal tract morphology (Ter Haar et al., 2021). Over time, some of these vocalisations show quite extreme convergence amongst pups. By adulthood, individual greater spear-nosed bats cannot reliably be identified by their foraging vocalisations (Knörnschild, 2014).

Between 2-10 weeks of age, greater sac-winged bat pups engage in exploratory vocalisation, producing chatters, screeches, whistles, trills, barks, fragments of territorial and mating calls of older bats, echolocation pulses, for up to a third of their waking time (Knörnschild et al., 2006; Knörnschild, 2014; Fernandez et al., 2021). From week three, pups begin to combine these into long, meaningless, context-independent, rhythmic sequences containing precise, adult-like sounds with immature, imprecise sounds. Over 75% of the sequences involve reduplication, whilst others comprise multiple different vocalisation types (Fernandez et al., 2021). These authors refer to these vocalisations as babble and theorise that producing these vocalisations furnishes developing bats with neuromuscular control that they later make use of when learning to vocalise like adult bats.

Both sexes produce babble, but females cease around 10 weeks (Knörnschild et al., 2006; Knörnschild, 2014). This apparently unnecessary vocal experience in females adds weight to the idea that babble is more exploratory than goal-directed. However, the vocal experience accumulated through babble may have the indirect evolutionary advantage of helping females to interpret males' territorial and mating calls more quickly and accurately later in life (Knörnschild, 2014; Fernandez et al., 2021). Male pups' babble gradually converges towards targets modelled by older 'tutor' males in a pup's immediate social circle (Ter Haar et al., 2021). By adulthood, male bats produce more acoustically similar vocalisations to those of conspecifics with whom they have stronger social ties than to those of genetically related individuals.

Babble in marmosets

Common and pygmy marmosets live in complex social groups comprising extended families and have complex systems for vocal communication (Elowson et al., 1998; Snowdon, 2017). From 1-10 weeks of age, marmosets produce rhythmic, repetitive vocalisations which contain a subset of the adult vocal repertoire produced with adult-like precision, and which have been termed babble or *gecker* (Omedes, 1985; Elowson et al., 1998; Snowdon, 2018). However, these infant vocalisations are produced context-independently and without meaning or reference. The frequency and distribution of vocalisation types in marmoset babble do not reflect the frequency and distribution of these sounds in adult vocalisations, suggesting that like human babies with Vocal Motor Schemes, infant marmosets also develop idiosyncratic production preferences.

Marmoset babble is primarily produced to elicit the attention of caregivers (Omedes, 1985). As with human babies, marmoset caregivers' responsiveness to babble may influence infants' motivation to repeat the behaviour. Pygmy marmosets have only been observed to respond silently to their infants' babble, picking them up and soothing them (Elowson et al., 1998). Common marmoset caregivers, however, *have* been found to respond with contingent vocal feedback, but only when an infant's babble happened to be at least mostly contextually appropriate (Takahashi et al., 2015). It is possible that the same may be true for pygmy marmosets though further study would be required to confirm whether this is the case. In common marmosets, contingent feedback gradually guides infants to produce more adult-like sounds in more appropriate contexts and reinforces and refines vocal turn-taking (Levinson, 2016; Snowdon, 2017). More voluble infant marmosets receive more contingent feedback and their vocalisations become adult-like earlier than less voluble infants (Snowdon, 2018).

Since adult marmosets show differential hemispheric involvement in prosodically and sequentially modulated vocalisations (Hook-Costigan & Rogers, 1998), it would be interesting to know whether infant marmosets show a similar differentiation in prosodically and sequentially modulated sounds in their babble,

Chimpanzees and bonobos

Whilst no rhythmic sequential behaviour approximating babble is seen amongst our closest cousins, some other behaviours which may be of interest to future laterality research have been identified (Oller et al., 2019). Infant chimpanzees' early grunts vary flexibly and systematically as a function of the affective state of the individual and it has been suggested that these acoustic variations, over time, form the foundations of later emerging contexts-specific calls (Dezecache et al., 2019). Infant bonobos produce voluntary exploratory sounds quite flexibly, across affective states and contexts (Oller et al., 2019). One unexpected finding of the present study is that human babies produce babble in a range of affective states, though it was previously suggested that babies only babble in a contented affective state (Vihman, 2014). Like some babies (Vihman et al., 1985), infant bonobos produce more of these exploratory sounds when they are alone than during social interactions suggesting that they are endogenously motivated (Oller et al., 2019; Ter Haar et al., 2021). While human babies' babble and pre-babble vocalisations are often met with contingent caregiver interaction, infant bonobos are more often soothed to silence by caregivers when they produce exploratory vocalisations (Oller et al., 2019). This lack of contingent feedback may go some way to explaining why adult bonobos vocalise much less than their infants (see Warlaumont et al., 2014 on the 'social feedback loop').

Summary

Vocal learning research weakens the position that language is a wholly *de novo* trait in humans (Crow, 2008; Balter, 2010). The fact that other animals acquire sophisticated and systematic communicative vocal behaviours through cultural transmission calls into question whether an innate language capacity is needed to bridge the gap between babble and language acquisition. Instead, the transition from babble into language may be understood as a process of vocal learning, driven by a domain-general capacity for social learning and an innate, survival-based urge to adapt our behaviour to express proximity to, and reinforce social bonds with, other members of our species and/or social group amongst species with extended childhoods (Carouso-Peck et al., 2021; Becker et al., 2022). In each species, genetics may define the parameters for the vocalisations available to the developing animal, but how each individual implicitly adjusts their own settings within those parameters may be a self-organising process guided by social experience and adaptive pressures.

This research also has a bearing on our understanding of what babble is and why we do it. Commonalities are observed between the characteristics and incidence of human and non-human babble and other early vocal behaviours. These commonalities support the idea that babble emerges endogenously through vocal exploration. Since non-human animals do not possess language, these commonalities also suggest that the neural architecture supporting babble is domain- and species-general rather than language- and human-specific.

The laterality of emotional processing

The next part of this chapter will consider the implications of findings from the present study regarding the laterality of positively valenced emotional facial expressions. In the present study, smiles and laughs were found to show slight right laterality (Model 1). Model 2 indicated no significant change in the laterality of smiles and laughs over developmental time ($p = 0.117$, Model 2). There are two major competing theoretical perspectives regarding the laterality of emotion, namely the Right Hemisphere hypothesis and Valence theory (Lindell et al., 2013). The former posits that emotional processing is predominantly supported by structures in the right hemisphere of the brain, while Valence theory, by contrast, posits that only negative emotional affect is right lateralised, whilst positive affect is handled by left lateralised networks (Lindell et al., 2013). For ethical reasons, only positively valenced emotion was analysed in the present study. If babies cried or fussed for more than a few moments, the video camera was turned off at the caregiver's and/or researcher's discretion. Nonetheless, the findings of the present study tentatively favour the Right Hemisphere Hypothesis. Although smiling and laughing showed only slight evidence of a left hemimouth bias with an estimated OAI of -0.009 (Model 1), this does not indicate *left* laterality for positively valenced expressions of emotion. Extending this study with a larger sample of participating babies and a longer data collection period could help to confirm whether the non-significant effect of persists and if so, whether it passes the threshold for significance.

Valence theory

Although Valence theory is not supported by the findings of the present study, research in this field may offer some insights of interest for interpreting this study's findings. Some brief background is provided for context.

Much of the evidence for Valence theory comes from adult human research in typical and atypical populations (e.g., Sackeim et al., 1982). In their 1982 review, Sackeim et al. comment on mood disturbances and the incidence of pathological laughter and crying in connection with right and left hemispheric damage respectively. The authors suggest that each hemisphere may be predisposed at the level of neurotransmitters towards handling either positive or negative emotion. However, Lindell (2013) points out that some types of right hemispheric damage can impair emotional processing regardless of valence and that treating individuals presenting with such impairments led to Hughlings-Jackson's (1874, 1915) earliest conceptualisation of the Right Hemisphere hypothesis.

Research with typical populations has also provided some behavioural and neuromuscular evidence supporting Valence theory (e.g., Schwartz et al., 1979). The left zygomatic major – one of the bilateral muscles controlling retraction of the cheeks and lip corners – shows greater movement than the right when adults talk about negatively valenced topics indicating right laterality, and the reverse is observed when adults discuss positively valenced topics, indicating left laterality (Schwartz et al., 1979). Furthermore, the magnitude of these asymmetries may reflect the strength of emotion experienced. The present study (Model 1, 2) and several previous orofacial asymmetry studies examining the movement of the lips have made findings to the contrary, suggesting that this asymmetry in muscle innervation may not be reflected in the overall shape of lip openings (Graves et al., 1982; Wylie & Goodale, 1988; Holowka & Petitto, 2002a).

In one facial composite study, adults were found to perceive left-left (L-L) chimeric images of posed negative emotional facial expressions (sadness, fear, anger, disgust) as more intense than R-R composites 73% of the time (Sackeim & Gur, 1978). The same adults judged L-L composites of positive emotional facial expressions to be more expressive than R-R composites only 45% of the time. Other adults have judged R-R composites of actors generating easy antonyms (e.g., for 'up' or 'fat') as more expressive than L-L composites (Schiff & MacDonald, 1990, p. 778). When actors were generating more difficult antonyms (e.g., for 'petulant' or 'acrimonious', p.778) L-L composites were perceived as more expressive. Actors reported negative feelings (frustration, anxiety) during the difficult task, and the authors suggest that this is reflected in their leftward orofacial asymmetry in this task. Another explanation may be that the easy and difficult tasks called upon differently lateralised linguistic operations. The capacity for word retrieval is predominantly left lateralised, but the right superior temporal gyrus supports comprehension when we encounter low frequency or unfamiliar words like those used here (Harpaz et al., 2009; Riès et al., 2016).

Only one study to date (to this author's knowledge) has found possible evidence for Valence theory in non-human primates. However, the strength of this interpretation is to be questioned. Common marmosets show leftward orofacial asymmetry for fear calls (*tsik*) and silent fear expressions indicating right laterality, whilst social contact calls (*twitter*) show left laterality (Hook-Costigan & Rogers, 1998). These social contact calls are positive affiliative signals. However, the authors note that *tsik* is produced by modulating pitch, whilst *twitter* involves sequential repetition and modulation the timing of open-close phase transitions. Hook-Costigan and Rogers argue that these findings may reflect activity in differently lateralised networks for producing melodic vs. sequential auditory information. Recall that pitch may be processed efficiently by the human right hemisphere, whilst producing and interpreting sequential vocalisations is more efficiently handled by the left hemisphere (Hook-Costigan & Rogers, 1998; Nishitani et al., 2005; Lindell, 2006; Minagawa-Kawai et al., 2011; Sussman, 2015; Riès et al., 2016; Hodgson et al., 2021).

Babies

Newborns show facial expressions of disgust and leftward asymmetry in activity in the parietal region (measured using EEG) when fed plan water and sweetened water elicits a stronger left hemispheric parietal response (Fox & Davidson, 1986). Citric acid water, designed to taste unpleasant, elicits a disgusted facial expression but a closely similar pattern of neural activity to sweetened water. Unlike the parietal regions, the frontal regions do not reliably show any comparable pattern of asymmetry in newborns (Fox & Davidson, 1986). By 10 months of age, more differentiated activity is seen in the frontal regions than in the parietal regions using EEG (Fox & Davidson, 1988). Greater left- than right-sided frontal activation is seen when 10-month-olds view images of actors posing happy expressions though the reverse is not see for images of actors posing negative emotion, meaning that these findings only partially support Valence theory (Fox & Davidson, 1988).

Valence aside, Fox and Davidson's (1986, 1988) findings, offer neuroimaging evidence of lateralisation and localisation of emotional processing as a function of developmental time during the first year of life. Fox and Davidson (1986) note that frontal areas involved in emotional processing and complex cognition in other non-human higher apes also show evidence of late and gradual functional integration. Model 2 indicates that any such structural and functional neural reorganisation was not reflected in the laterality of babies' smiles and laughs between 5-12 months of age. However, Fox and Davidson's (1986, 1988) neuroimaging evidence of changes in cortical organisation provides motivation for further closer analysis of possible age-related changes in the laterality of emotional facial expressions in future. Fox and Davidson (1986) also comment on the apparent dissociation between facial and neural responses, suggesting that contextual factors and the dynamicity of emotional facial expressions may sometimes give rise to mismatch between neural activity and observable orofacial behaviour. Incidental dissociation between neural and facial activity during the experiencing of an emotion may account for the lack of a significant interaction between smiles and laughs and age, as well as for the variability seen in babies OAIs for smiles and laughs in the present study (Model 1, 2; Figure 9, Figure 10).

The Right Hemisphere hypothesis

The Right Hemisphere hypothesis has received more support from the wider literature and is supported by the findings of the present study.

Adults

As with language (see Chapter 3), neuroimaging research with adults has identified a distributed but predominantly right lateralised task-dependent networks of brain areas involved in emotional processing (Häberling et al., 2016; Häberling & Corballis, 2016). These areas include the right inferior frontal gyrus and right superior temporal sulcus; the bilateral inferior occipital gyri, fusiform gyri, ventromedial prefrontal cortices and amygdalae, and portions of the left cerebellum (Sato et al., 2019). The right hemisphere's proficiency at processing global, holistic visuospatial and auditory information may make these parts of the brain particularly well-suited to analysing and responding to movements of the face and changes in tone of voice (Sussman, 2015) – both of which contain information which make be useful for survival or social bonding like the emotion being experienced, its valence, and its urgency.

Likewise, behavioural research has found greater left hemimouth activity during expressions of both positive and negative emotion indicating right laterality (Borod et al., 1983, 1988; Wylie & Goodale, 1988). Adults' lips show greater displacement in the left hemimouth than the right during spontaneous, but not posed positive emotional facial expressions (Wylie & Goodale, 1988). Video data of posed and spontaneous expressions of happiness and sadness rated for asymmetry, intensity, and valence, has revealed some differential effects of sex but not of valence (Borod et al., 1983). Negative facial expressions showed greater intensity and expressivity in the left hemiface for both sexes. However, only men showed significantly greater left than right hemifacial intensity and expressivity for positively valenced emotional facial expressions. Women showed a weaker laterality in positive facial expressions than in negative ones (Borod et al., 1983). Men have shown significantly greater left hemifacial intensity of expression in posed facial expressions of five negative emotions (sadness, anger, fear, disgust, confusion) and three positive emotions (happiness, positive surprise, and sexual arousal) (Borod et al., 1988). (No female actors took part in this study). The present study set out to explore the laterality of babble, using the laterality of emotion as a point of contrast. As a result, sex differences in the laterality of emotional facial expressions were not explored in Models 1-7. As mentioned earlier in this chapter, the sample of participants involved in this study was unbalanced, containing 5 male babies and 3 female babies. The present study could be extended by adding data from more female babies could be added to the present corpus and constructing models to test whether babies show any differences in the stability, direction, or strength of their laterality with regard to emotional processing associated with sex.

Some facial composite studies have made complementary findings to Borod and colleagues (1983, 1988). Unlike Sackeim and Gur (1978), Asthana and Mandal (1998) and Nagy (2012) found L-L composites of adults posing happiness to be perceived as more intense than R-R composites. L-L composites of posed and spontaneous expressions of grief show greater intensity of emotion than R-R composites, indicating right laterality (Cacioppo & Petty, 1981). Posed and spontaneous R-R composites of facial expressions associated with thoughtfulness and mathematical cognition show greater intensity than L-L composites, indicating left laterality (Cacioppo & Petty, 1981). Cacioppo and Petty (1981) tentatively suggest that their cognitive conditions (posing thoughtfully and reading a mathematics textbook aloud), may have elicited positive emotional responses and that their findings support to Valence theory. An alternative, less speculative possibility is that their cognitive conditions activated predominantly left lateralised task-dependent networks involved in voluntarily controlled sequential orofacial movement, speech, reading, and complex mathematical cognition (Graves et al., 1982; Wyler et al., 1987; Amalric & Dehaene, 2016; Häberling & Corballis, 2016; Häberling et al., 2016; Hodgson et al., 2021).

Babble, like reading aloud, involves using the lips, tongue, and jaw, which are also involved in speech (see also Wolf & Goodale, 1987; Gannon et al., 1998; Salmelin & Sams, 2002; Müller & Basho, 2004; Nishitani et al., 2005; Saarinen et al., 2005; Hodgson et al., 2021). In the present study, babble showed significant left laterality with an estimated OAI of 0.026 (Model 1). Vihman (2014) has stated that babble is produced when babies are in a contented affective state. Considering the central claim of Valence theory, it could potentially be argued that its left laterality reflects some underlying positive affective neural activity. This is unlikely for two reasons. Firstly, one unexpected observation made during the present study was that babies babbled both when contented and when fussing and sometimes when crying. Secondly OAIs for smiles and laughs – unarguably expressions of positive emotion – showed significantly different laterality to babble. Whilst babble showed left hemisphere dominance, smiles and laughs – like the positive and negative emotions examined by Borod and colleagues (1983, 1988) and by Cacioppo and Petty (1981) – showed slight right hemisphere dominance, with a category estimate of -0.009 (Model 1). This implies that babble and positive facial expressions were not underlain by the same neural architecture as one another in babies in this study.

Future research could explore whether contented and fussy babble show any difference in laterality. Both fussy and contented babble utterances were included in the present OAI corpus, though fussy babble utterances were too few in number to explore this idea statistically. Collecting data of babies experiencing negative emotion poses ethical concerns, particularly where repeated measures are required for analysis. This author would therefore recommend that any study intending to explore a possible between-valence difference in the laterality of babble should strongly consider adopting a cross-sectional design as Best and Queen (1989) and Schuetze and Reid (2005) have done in order to limit the discomfort and distress of each baby, their caregiver/s, and the researcher.

Non-human animals

Behavioural research with some non-human animals has also found evidence favouring the Right Hemisphere hypothesis and suggesting that right laterality for emotional processing may date back at least as a far as a common haplorrhine ancestor[35] (c.27.0 - 30.6 MYA) but perhaps even further (Siniscalchi et al., 2008; Lindell, 2013; Siniscalchi et al., 2016; Siniscalchi et al., 2018; Kumar et al., 2022). Some of this evidence comes from orofacial asymmetry research. For example, species-typical emotional vocalisations and facial expressions in chimpanzees show right hemisphere dominance regardless of emotional valence (Fernández-Carriba et al., 2002a, 2002b; Losin et al., 2008; Wallez et al., 2012). Fernandez-Carriba et al. (2002a, 2002b) examined the laterality of aggressive vocalisations (*pant-hoot*), and facial expressions of distress (*silent pout*), fear/anxiety (*silent bared-teeth display*), protest (*staring bared-teeth scream face*), and positive affiliation (*play face*). Hemimouth measurement was undertaken, and a facial composite experiment was conducted in which humans rated L-L and R-R composites of chimpanzees faces for intensity of emotion. In both sub-studies, all expression types showed right laterality. Interestingly, when *chimpanzees* are presented with neutral-happy facial composites of *humans* and requested to pick the happiest looking image, they select the image wherein the smiling hemiface is in the left visual field and the neutral hemiface is in the right 62% of the time (Morris & Hopkins, 1993). Humans make the same selection 63% of the time (Morris & Hopkins, 1993). Facial emotion processing is right lateralised in both species. Losin et al.

---

[35] In grey mouse lemurs, a strepsirrhine primate, males show strong *left* laterality for perceiving negatively valenced species typical emotional calls, counter to the predictions of both Valence theory and the Right Hemisphere hypothesis.

(2008) examined the laterality of aggressive vocalisations (*pant-hoot*) and excited affiliative vocalisations (*food-bark*), alongside two artificially learned, communicative vocalisation types. Learned vocalisations showed left hemisphere dominance (see Chapter 3), but species typical emotional vocalisations of both valences showed right hemisphere dominance (Losin et al., 2008). Wallez et al. (2012) extended Losin et al.'s study with a separate population of chimpanzees and confirmed these findings.

Both humans and chimpanzees possess 23 facial muscles, giving us a more closely similar range of facial mobility to one another than to other primates and non-primates who have fewer and differently configured facial muscles (Diogo et al., 2009; Lindell, 2013). This may explain why cross-species similarities in emotional processing can be observed. However, human facial expressions show finer and more nuanced motor control than those of chimpanzees, perhaps reflecting differences in the range of emotional expressions required for survival and social cohesion in these species (Wallez & Vauclair, 2012; Lindell, 2013).

Despite differences in facial physiology (Diogo et al., 2009), baboons and macaques also show evidence of right hemisphere dominance for producing emotional facial expressions Wallez and Vauclair (2011, 2012) found evidence of right lateralised emotional processing in two species of Old World monkey – baboons and macaques. Adult and adolescent olive baboons show strong right laterality in screeching – an agonistic vocalisation type – though not in agonistic eyebrow raising (Wallez & Vauclair, 2011). Recall that neuromuscular control of the mouth is cross-lateral, whilst control of the upper facial muscles is bilateral (Rinn, 1984; Wallez & Vauclair, 2011). Lip-smacking and copulation calls – both affiliative vocalisation types – show no clear and consistent hemispheric dominance, though the Wallez and Vauclair (2011) link this to the low still-frame sample sizes of 70 and 24 respectively compared with the still-frame sample of screech (n = 116), and the marked variability in the ages of the baboons from whom these samples were drawn (p. 167). Baboons show a preference for keeping social partners in the left visual field during negatively valenced encounters and Mangabeys – another species of Old World monkey – favour the left visual field during affiliative interaction (Lindell, 2013). Pacific walrus (from whom we diverged around 89.6 - 97.4 MYA, Kumar et al., 2022) and Indian flying foxes (divergence at 424.2 - 440.0 MYA, Kumar et al., 2022) likewise demonstrate a left visual field preference during face-to-face interaction with conspecifics, suggesting that some right hemispheric architecture involved in emotional processing pre-dates primates (Giljov et al., 2018). Further, this preference is observed in some infant and adult mammals alike (Giljov et al., 2018). A body of research not discussed here has explored the laterality of emotion in non-primate animals (Siniscalchi et al., 2008; Lindell, 2013; Siniscalchi et al., 2016; Siniscalchi et al., 2018) and some of this research, too, has found evidence favouring the Right Hemisphere hypothesis.

The overlapping findings of this human and non-human research tentatively suggest that the laterality of emotional processing, like the laterality of some language-related functions, may be partly genetically determined. However, infant research suggests that some experience-dependent hemispheric specialisation also takes place during development.

Babies, toddlers, and children

Like Graves et al. (1982), Holowka and Petitto (2002a) found babies' smiles show a significant tendency towards right laterality. Laterality research using other behavioural techniques with babies, toddlers, and children has made less straightforward and consistent findings with regard to the stability and direction of orofacial asymmetry during emotional facial expressions and emotional processing (e.g., Fox & Davidson, 1986, 1988; Best & Queen, 1989; Rothbart et al., 1989; Schuetze & Reid, 2005; Workman et al., 2006; Nagy, 2012; Lindell et al., 2017). Where these other studies agree is in the idea that some process of post-natal rightward lateralisation takes place gradually over several years.

Facial composite studies have made equivocal findings (Best & Queen, 1989; Rothbart et al., 1989; Nagy 2012). R-R chimeric images of 7-13-month-old babies producing smiles and distressed expressions show greater expressivity than L-L composites indicating *left* laterality rather than right laterality (Best & Queen, 1989). In a hemifacial rating study with babies aged 3, 6.5, 10, and 13 months, facial expressions of happiness and distress showed consistent left hemisphere dominance from the early months of life (Rothbart et al., 1989). By contrast, another facial composite study found no consistent evidence of orofacial asymmetry in either direction in the smiles of 0-12-month-olds or 3-8-year-olds, (Nagy, 2012). Eighteen-to-thirty-two-year-olds in this same study indicate right laterality in their smiles. The high real-time variability in the laterality of babies' smiles and laughs in the present study may shed some light here. In the present study, 540 still-frames of smiles and laughs were taken from eight babies. Best and Queen's study used images of 10 infants in total, 4 showing distress and 6 showing happiness, and Nagy's study used one photograph per baby (n = 135) and child (n = 78). Rothbart et al.'s study involved 59 babies, but the number of images rated is not stated. Having collected a large corpus of laterality data containing repeated measures from each baby, the present study may have captured more representative patterns of individual behaviour that were less susceptible to interference from between-individual and within-individual real-time variability. Having used mixed effects models, designed to control for random individual variability, the present study may also have been able to identify subtler effects of laterality. Additionally, using hemimouth area measurement may have allowed the present study to identify smaller, less easily observable differences in the size and shape of lip openings than rating studies, which rely on visual inspection.

Studies using hemimouth area measurement with human and non-human infants have made less equivocal findings (Fernandez-Carriba et al., 2002a, 2002b; Holowka & Petitto, 2002a; Schuetze & Reid, 2005; Wallez & Vauclair, 2012). Between them, these studies have found both positive and negative emotional facial expressions and vocalisations to show right hemisphere dominance. Findings from the present study are more consistent with findings from these studies. This adds weight to the argument that hemimouth area measurement offers a more objective, veridical, and replicable way to analysed orofacial asymmetry than other methods (Fernandez-Carriba et al., 2002a, 2002b). Two distress vocalisation types in infant baboons (*geck* and *moan*) and one type in infant macaques (*coo*) – have been found to show significant right hemisphere dominance (Wallez & Vauclair, 2012). Another macaque distress vocalisation – *scream* – shows left hemisphere dominance, counter to the predictions of both Valence theory and the Right hemisphere hypothesis. Infant macaques and baboons scream and geck (respectively) when a request to breastfeed is denied by the mother, and they produce coos and moans (respectively) to self soothe and to seek attention from other potential caregivers (Wallez & Vauclair, 2012). Human 1-2-year-olds show greater right-sided lip opening in negatively valenced emotional facial expressions, increasing in strength throughout the second year of life (Schuetze & Reid, 2005). This cross-sectional study found no reliable evidence of any asymmetry for positive expressions in 1-2-year-olds.

The estimated OAI of -0.009 for smiles and laughs in the present study (Model 1) indicates right laterality, but also demonstrates weaker laterality than babble (0.026). These findings may be more closely aligned with those of Schuetze and Reid (2005) than with those of Holowka and Petitto (2002a). Like Schuetze and Reid's study, the present study also found no statistically significant change in the laterality of positively valenced emotional facial expressions over time (Model 2). The authors assessed change over time by comparing groups of 12-, 18-, and 24-month-olds to one another. A minimum of one still-frame for positive emotion and one for negative emotion for each of 41 babies and toddlers was analysed (n = ≥82, though the exact quantity and distribution of still-frames analysed is not stated). In the present study, 8 babies were compared against themselves and each other, with repeated measures taken both within and across visits, resulting in 37-99 still-frames per baby (mean = 68). This was undertaken to capture individual variability in both real- and developmental time. Because only a non-significant effect of age was detected in the present study, it is not possible to conclude that orofacial asymmetry analysis shows evidence of increasing right lateralisation in smiles and laughs during the first year of life. Nor is it possible to conclusively exclude the possibility that such an effect exists in the wider population. It may be that the right laterality seen in adults' smiles emerges sometime after 24 months.

Indeed, behavioural testing with toddlers and children demonstrates emotional functions are not fully right lateralised until much later in development. For example, when asked to select the happiest looking of two photographs, 3-7-year olds consistently select images in which the subject of the photograph has the left cheek oriented towards the camera over photographs wherein the right cheek is oriented towards the camera (Lindell et al., 2017). When asked to smile for photographs themselves, however, adults but not children will reliably orient their own left cheek towards the camera. Ten-to-eleven-year-olds show a significantly stronger left visual field preference than 5-6-year-olds when the judging the intensity of emotional-neutral facial composites when then emotion pictured is sadness, fear, anger, disgust, or surprise (Workman et al., 2006). When the emotion pictured is happiness, children's performance is more similar. These findings demonstrate right hemispheric specialisation for emotional processing persisting into mid to late childhood. Furthermore, this increasing lateralisation is positively correlated with measures of precision and accuracy in understanding the emotions of others using information from a person's eyes and facial expressions, and external dynamic contextual factors (Workman et al., 2006). Emotional *expression* may come to be right lateralised because it builds on foundations laid by gradual strengthening of right hemispheric networks for emotion *perception* (Tierney & Nelson, 2009). With reference to other primates, Lindell (2013, p. 7) writes that "[o]bservation of others' emotional expressions may increase the magnitude of the hemispheric asymmetry for emotional processing in diurnal primates via experience-dependent processes." (p.7). in light of these findings, hemimouth area measurement could be used longitudinally with 3-11-year-olds to test if and when the experience-dependent changes outlined here are associated with any change in the orofacial asymmetry of emotional facial expressions.

Alternative views

In response to the debate regarding the underlying nature of emotion, some authors have put forward alternative conceptualisations of emotional processing which seek to unite the more well-supported aspects of Right Hemisphere hypothesis and Valence theory in a complementary fashion (Killgore & Yurgelun-Todd, 2007; Ross, 2021).

With regard to emotion perception, (Killgore & Yurgelun-Todd, 2007) have suggested that the evidence supporting Right Hemisphere hypothesis and Valence theory may highlight the workings of "two interrelated systems" (p. 247). These include a more localised system in the posterior portions of the right hemisphere, adept at processing any emotion of either valence, and a weaker and more dispersed system in the left hemisphere. The right hemisphere system is suggested to be more differentiated, resulting in a right hemispheric advantage for processing negative emotions, which are more numerous and more nuanced. The left hemispheric system, by contrast, is suggested to be less differentiated, making it reasonably well-suited to processing positive emotions, which are fewer in number (and perhaps requiring less urgent responses), but much less efficient at processing negative emotion. Unlike the right hemispheric system, the left hemispheric system may need to recruit support from a more dispersed network of left hemispheric neuronal clusters in the left orbitofrontal cortex and ventral striatum when handling negatively valenced emotions (Killgore & Yurgelun-Todd, 2007).

In the data collected for the present study, sometimes babies' smiles and laughs appeared unprompted, but some were responses to perceiving positive emotion in caregivers and the researcher. Since neural networks in the brains of developing babies are more dispersed, degenerate, and redundant than those of older children or adults (Thelen & Smith, 1994; Mareschal et al., 2007), the availability of multiple different networks capable of processing positively valenced emotion may explain some of the high real-time variability seen in babies' OAIs for smile/laugh. The network or hemisphere 'selected' to process positive emotion in one instance may not be the same network or hemisphere that carries out this work in another instance. Which hemisphere is activated at any given moment in time may be determined by dynamic contextual factors relating to the other actions or cognitive processes that the baby's brain is carrying out around the time of that moment.

With regard to both perception and production, Ross (2021) posits that emotions may be organised into two categories – primary and social – each requiring different types of processing. Primary emotions are suggested to be innate, predominantly (though not exclusively) negative, and to contribute directly to decisions supporting survival. For instance, disgust at bitter flavours may serve to protect the individual from consuming or handling substances that may be poisonous or otherwise inedible. Social emotions are acquired through experience and are those that support within-group social cohesion. These includes things like culture-specific ways of handling rejection, saving face, or signalling gratitude or affection. Whilst the facial expressions associated with the six basic emotions are observed to be universal (Ekman, 1971, 1992), voluntarily modified facial expressions associated with more nuanced social behaviours may be specific to the groups that use them (Ross, 2021). Ross (2021) suggests that perception and production of primary emotions may be a predominantly right hemispheric job, whilst perceiving and producing socially learned, voluntarily modulated orofacial movements involved in social emotional facial expressions may involve left hemispheric structures to a greater degree than the primary emotions. As mentioned above sometimes babies in the present study smiled and laughed apparently unprompted by others, but some smiles and laughs were produced in the context of a social interaction. The laterality of a smile or laugh may, then, influenced by the context in which it is produced. This, too, may account for some of the rather high variability observed in babies' OAIs for smiles and laughs.

Conclusions

Very little research to date has examined the laterality of babble, though exploration of both phenomena – laterality and babble – has potential to provide insights with clinical and academic significance.

The present study has made several novel contributions to research in the fields of laterality and pre-linguistic vocal and emotional development, which may also have implications for our understanding of the ontogeny and phylogeny of language. This study has established that Holowka and Petitto's (2002a) finding that babble is left lateralised can be replicated using hemimouth area measurement. By replicating the findings of earlier laterality research using an updated, rigorous, and objective method, the present study has validated the usefulness of this method for laterality research with babies. Hemimouth area measurement has significant advantages over neuroimaging and neurosurgical methods, being cost-effective, sustainable over longitudinal data collection periods, and minimally invasive, optimising the ecological validity of any resulting findings. This method also has clear advantages over Graves et al.'s (1982) method for analysing orofacial asymmetry, used by Holowka and Petitto (2002a), in that it provides data about the strength and consistency as well as the direction of lateralisation in an orofacial gesture.

The present study has also made novel findings using this updated method for hemimouth area measurement. Firstly, having explored not only the directions of babies' lateralisation for babble but also the strength and stability of this lateralisation, this study has established that babble shows considerable variability in hemispheric involvement in real time and significant left lateralisation over developmental time. Secondly, by observing lateralisation in babble longitudinally, the present study has established that babble is not left lateralised but rather is *right* lateralised at the time when it first emerges. These findings have important implications for the Linguistic and Motoric hypotheses alike. Proponents of both of these theories have justified their stances based on the assumption that the syllables of babble emerge well-formed and mature-sounding. The findings reported here contribute novel evidence to support the argument put forward by (1980), who illustrated how mature-sounding babble emerges, not *de novo*, but rather out of much less mature-sounding sounds. The present study provides evidence that babble is also not adult-like in terms of the distribution of neural activity that underlies it. Thus,the present study has shown that that babble is not stably or consistently left lateralised like adult speech and language unlike speech and language in adults. This calls into question the idea that babble is underlain by an innate left hemisphere brain-based capacity for language. Rather, it seems more likely that the stable lateralisation seen in adults comes about as a function of developmental time and articulatory experience.

The extent of the variability in the laterality of babble over real and developmental time observed in this study may reflect the higher redundancy and degeneracy of babies' cerebral organisation and the gradual, experience-dependent processes of hemispheric and regional specialisation that characterise the transition from infancy to childhood to adulthood. The finding that producing bisyllabic babble utterances is associated with significant left lateralisation over developmental time adds weight to the idea that hemispheric specialisation is influenced by productive (as well as perceptual) experience and illustrates which experiences may be most influential in this process during the first year of life. Though recall that no significant leftward shift was observed for polysyllabic babble, perhaps because polysyllabic babble was more left lateralised than bisyllabic babble when it first emerged.

The primary focus of the present study was the laterality of babble. However, this study aldo made one finding of interest regarding lateralisation for positive emotional facial expressions. That smiles and laughs showed only slight right laterality is neither consistent with Valence theory, nor wholly consistent with the Right Hemisphere hypothesis. However, this finding does align with some previous infant research in suggesting that the laterality of positively valenced emotional facial expressions is also subject to some experience-dependent process of lateralisation, persisting throughout infancy and childhood.

Finally, this thesis has put forward an alternative hypothesis concerning the underlying nature of babble – the Old Parts, New Machine hypothesis. The Old Parts, New Machine hypothesis unites observable phenomena in development and combines these with insights from Dynamic Systems theory in formulating a biologically and psychologically plausible explanation for what babble is and why we do it. Throughout this thesis, it has been argued that this novel theory of babble can offer better account than previous competing theories of babble – the Linguistic and Motoric hypotheses – for how babble emerges endogenously out of in-repertoire, domain-general behaviours and capacities, how and why it changes in form and laterality over time, and how it may influence and constrain the transition into language acquisition.

It is hoped that the present study has also laid foundations for future research in this and related fields. Throughout this thesis, possible directions for future research have been considered. Three main directions emerge. Firstly, the under-researched field of laterality in pre-linguistic development and first language acquisition could be expanded by research examining the laterality of marginal babble and other early vocalisation, of first words, and of more and less stable consonants. Secondly, the method used here could be tested in clinical settings. If babble is atypically lateralised in atypically developing babies and children, then having this knowledge could help us to identify those in need of additional support early in development and could inform our understanding of how to construct and implement effective support. Finally, there are as yet unexplored avenues in comparative laterality research. In particular, the relationship between laterality and vocal learning across species in which it is observed may hold valuable insights into how and why we babble and do language. In particular, there is a strong case for using hemimouth area analysis in comparatively laterality research with humans and non-human primates.

Limitations

Owing to its intensive nature, this study involved a fairly small sample of participants (N=8). As a result, the generalisability of the findings reported here may be limited. Firstly, with the exception of one family, the majority of participants in the present study were mid-SES families in which one or more caregiver/s held one or more university degree/s. Social, psychological, and developmental research involving voluntary participation often disproportionately attracts mid- and high-SES families, resulting in underrepresentation of lower SES groups and individuals (Baucom et al., 2018). Strategies for improving accessibility and representativity in research include things like recruiting through healthcare providers in areas with a higher density of marginalised individuals, fostering strong, positive relationships with a community some time prior to data collection, and tailoring the proposed term of commitment and the compensation plan to the needs and interests of the target population (Baucom et al., 2018). In the present study, recruitment was conducted through the York BabyLab because this was known to the researcher to be a reliable source for participants in the York area. The York area was selected for logistical reasons. Being at the start of their career, this author had limited opportunities prior to recruitment to engage with the public about the actual and expected impact of their work. Participants were compensated for their time, though the sum allotted was modest, owing to the limited availability of funds for participant payment in doctoral projects. Whilst every effort was undertaken to ensure that the data collection programme worked around participating families' schedules, the overall term of commitment in the present study was somewhat inflexible. This decision was undertaken, with the consent of participating families, to ensure that the data collected was comparable both across babies and with the data collected by Holowka and Petitto (2002a). Secondly, to participate in this study, babies must have been born at full term and must have been typically developing, with no known hearing impairment. Babies with, or at risk of, diagnoses of developmental delays or disabilities may show different patterns in their developmental trajectories or outcomes with regard to lateralisation. The decision to include only typically developing babies in this study was undertaken because very little research to date has examined orofacial asymmetry in pre-linguistic vocal behaviour. The present study aimed to begin to address this gap in the research by first examining laterality in typically developing babies. It is the hope of this author that the present study might pave the way for future work examining laterality in pre-linguistic vocal behaviour in typically and atypically developing children by providing a point of comparison.

A key limitation of oral asymmetry analysis is that, whilst it can tell us about the strength of laterality and can trace lateralisation over time, it cannot tell us which specific areas within each hemisphere are active during orofacial behaviours of interest. Nonetheless, its notable advantages as a rigorous but non-invasive method for analysing laterality mean that it can be applied across ages and species in naturalistic and experimental data. This can allow us to draw direct comparisons between these groups where similar methods for analysing orofacial asymmetry are used.

Naturalistic orofacial asymmetry analysis allows us to build a more fine-grained and ecologically valid picture of participants' laterality in behaviours of interest than experimental research can. However, its affordances can also make it vulnerable to certain constraints. Unbalanced corpora are characteristic of orofacial asymmetry research. Table 32 below details the sample of still-frames or photographs deemed analysable in some of the experimental and naturalistic studies which were reviewed in this thesis, and which informed the design of the present study. Some oral gestures that involve lip retraction or open-lipped onsets are more difficult to capture on film and analyse (Wolf & Goodale, 1987). Where research uses naturalistic data, participants' freedom to move around may make capturing behaviours of interest more challenging (Fernández-Carriba et al., 2002a, 2002b). Additionally, still-frame extraction, preparation, and measurement is particularly time-consuming (Wolf & Goodale, 1987), and this imposes constraints on the volume of data that can be analysed. In the present study, OAIs were generated for a corpus of 1494 still-frames from 8 babies aged between 5-12 months. This is a relatively high sample size for orofacial asymmetry analysis. However, within this corpus, some babies produced larger or smaller proportions of each subtype of babble (see Table 8). Some babies produced more monosyllabic babble whilst other favoured bisyllabic or polysyllabic babble (see Table 9). Most significantly, for some babies, no data at all was available either before or after the emergence of their first Vocal Motor Scheme.

Finally, ways in which the present study could be extended have been identified throughout this thesis. Comparing marginal babble to other subtypes of babble could provide more insight into the relationship between articulatory experience and laterality. Owing to the design of the data collection programme, too little marginal babble was captured to be included in analysis. The present study was designed to focus on canonical babble (singleton, reduplicated and variegated babble). To limit the rather intensive term of commitment for participating families, it was decided that, after the two baseline sessions at 5 months, families should only resume participation once their babies started to produce canonical syllables. Comparing first words to babble could yield information pertinent to the debate concerning the underlying nature of babble. The age at which the first words emerge is highly variable and collecting data at twice-monthly intervals from babble emergence to first word emergence could have resulted in very different terms of commitment for each family. This would raise ethical issues relating to inequality in the imposition on families' time and in the proportional value of the compensation offered. Some further analyses could also be conducted using the existing data gathered in this study. The actual distribution of babble utterances of each length and each subtype in the audio-visual recordings could be calculated and analyses performed to test whether predictions about order of emergence are borne out by the babies in this study. The relationship between the laterality of babble and the number of consonants a baby has in repertoire at the time of each home-visit could be investigated to test for an effect of articulatory skill differentiation. The laterality of babble utterances containing Vocal Motor Scheme consonants and non-Vocal Motor Scheme consonants could be compared to test whether well-practiced, stable behaviours show stronger or less variable laterality than less stable behaviours. These ideas fell outside the scope and timescale of the present study, but it is hope that they will be explored further in future.

Study	*Still-frame quota*	*Method*
Graves et al. (1982)	Participants: 196 typically developing adults aged 18-43 years, 98m, 98f Experiment 1: n=630 from 63 participants; Experiment 3: n=462 from 42 participants TOTAL = 1092	lip angle measurement
Wyler et al. (1987)	Participants: 63 typically developing adults, mean aged 20-43 years (quantity and distribution of still-frames not detailed)	lip angle measurement
Wolf and Goodale (1987)	Participants: 24 typically developing adults aged 19-27 years, 12f, 12m Meaningless syllable production: ma: n=24 from 24 participants	lip displacement measurement

Study	*Still-frame quota*	*Method*
	pi: n=24 from 24 participants ba: n=23 from 23/24 participants Silent non-speech mouth movement lip open: n=14 from 14/24 participants lip retract: (quota not stated, high attrition mentioned) blow: n=0 from 0/24 participants TOTAL = 85 (excl. retract)	
Wylie and Goodale (1988)	Participants: 35 typically developing adults, aged 19-30 years, 19f, 16m (actual quotas and distribution note stated. Spontaneous smiles: min. n=32 from 32/35 participants; posed smiles min. n=27 from 27/35 participants. Min. TOTAL = 69)	lip displacement measurement
Hook-Costigan and Rogers (1998)	Participants: 11 juvenile and adult marmosets, 7f, 5m Silent fear expression: min. n=110 from 11/11 marmosets vocalised fear expression ’tsik’: min. n=80 from 8/11 marmosets) social contact call ‘twitter’: min. n=80 from a partially different group of 8/11 marmosets). Statistical analyses performed on data from 6/11 individuals. Min. TOTAL = 180	hemimouth area measurement
Hausmann et al. (1998)	Participants: 39 students aged 19-42, 20f, 19m Discrete condition: syllable:n=39 from 39/39 participants non-word: n=39 neutral word: n=39 negatively valenced word: n=39 Serial condition: syllable: n=117 from 39/39 participants non-word: n=117	lip height measurement

Study	*Still-frame quota*	*Method*
	neutral word: n=117 negatively valenced word: n=117 TOTAL = 624	
Holowka and Petitto (2002a, 2002b)	Participants: 10 5-12 month-olds, 5 French, 5 English TOTAL = 150 (actual distribution of data not detailed. Assuming a balanced corpus, this may mean: babble (n=50 from 10 babies); non-babble (n=50 from 10 babies); smile/laugh (n=50 from 10 babies).	lip angle measurement
Fernández-Carriba et al. (2002a, 2002b)	Participants: 36 juvenile and adult chimpanzees hoot: n=39 from 22 chimpanzees play face: n=29 from 18 chimpanzees pout: n=14 from 9 chimpanzees silent bared-teeth scream face: n=31 from 20 chimpanzees vocalised scream face: n=19 from 11 chimpanzees neutral face: n=51 from 30 chimpanzees TOTAL = 183	hemimouth area measurement
Schuetze and Reid (2005)	Participants: 41 infants and toddlers aged 12-12 months, 23f, 18m (actual quota and distribution not reported). positively valenced emotional facial expression: min n=41 negatively valenced emotional facial expression: min n=41 Min TOTAL = 82	hemimouth area measurement
Losin et al. (2008)	Participants: 69 adult and juvenile chimpanzees pant-hoot: n= 68 from 32 chimpanzees food-bark: n=73 from 68 chimpanzees raspberry: n=100 from 35 chimpanzees extended grunt: n=26 from 10 chimpanzees TOTAL = 267	hemimouth area measurement
Wallez et al. (2012)	Participants: 42 + 69 adult and juvenile chimpanzees pant-hoot: n=47 from 23 chimpanzees	hemimouth area

Study	*Still-frame quota*	*Method*
	food-bark: n=87 from 11 chimpanzees raspberry: n=219 from 21 chimpanzees extended grunt: n=11 from 4 chimpanzees kiss: n=67 from 11 chimpanzees TOTAL = 431 Combined with sample from Losin et al. (2008) above: OVERALL TOTAL 698	measurement
Wallez and Vauclair (2012)	Baboons: 20 infant baboons distress call: n=85 from 17 baboons moan: n=29 from 11 baboons geck: n=56 from12 baboons TOTAL 230 Macaques: 11 infant macaques coo; n=77 from 11 macaques scream: n=68 from 9 macaques TOTAL: 145 OVERALL TOTAL: 375	hemimouth area measurement

Table 32 Photograph or still-frame sample sizes of previous orofacial asymmetry studies.

Appendices

Appendix I: Information Sheet

DEPARTMENT OF
LANGUAGE AND
LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Phone number 07956 052 247
Email fo507@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 project: Laterality in Babble

Researcher: Florence Oxley

What is the research about?

I am interested in finding out about how babble contributes to brain development and vice versa.

Who is carrying out the research?

The research is carried out by Florence Oxley, a PhD student with the Department of Language and Linguistic Science, under the supervision of Dr. Tamar Keren-Portnoy and Dr. Eytan Zweig.

Who can participate?

I am looking for families with healthy, normally developing babies born at full term, who are aged 5 months or under, and who have not yet started to babble.

What does the study involve?

I will visit you at home between 13-20 times over 6-7 months (usually 12 times). Each visit will last around 60-90 minutes (usually less). I will ask you to play with your baby as you normally would for 30 minutes, and I will video and audio record this. You will be with your baby throughout each session.

I will also provide you with a Developmental Diary, which I will ask you to update as and when your baby starts to show new behaviours. This information will be sent to me. I will run through the diary with you during the first visit. During the final session of the study I will provide you with a Communicative Development Inventory to record any words that your baby might understand or say. Again, I will run through this with you during the session when you receive it. [REVISED - REMOVED]

Finally, I will ask you about your education. You do not need to tell me this information if you do not want to.

You will be paid £10 and given a BabyLab t-shirt for your baby at the first visit and a further £50 upon completion of the study (£60 in total). If you withdraw before completing the study, you may keep the £10 and the BabyLab t-shirt, and you will receive a sum of the £50 completion fee, proportional to the term of your involvement (see table below, where 'Full participation' assumes ending the study when your baby is 12 months of age. Babies develop at different rates, so some families might only complete around 12 sessions before their baby is 12 months, and still be considered to have participated fully.)

Sessions completed	**4 sessions**	**8 sessions**	**12 sessions**	**Full participation**
Fee paid	£10 (+£10 and t-shirt)	£25 (+£10 and t-shirt)	£35 (+£10 and t-shirt)	£50 (+£10 and t-shirt)

Do I have to take part?

You do not have to take part in the study. If you do decide to 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 take part you will still be free to withdraw before the end of the last visit without giving a reason, even in the middle of a session. If you withdraw from the study, I will destroy your data and will not use it in any way.

What are the possible risks of taking part?

It is possible that your baby will fuss or be unhappy during the visit. If this happens, I will be happy to postpone or reschedule visits to accommodate you or your baby. If you feel your baby is distressed at any time, you are free to stop the session. We do not think there is any risk to you or your child from being video and audio recorded. You can also ask for the video-recording from a certain session to be deleted at the end of that session if you are unhappy about it being retained by the University.

Are there any benefits to participating?

Participants will be provided with a copy of their baby's Developmental Diary as a souvenir. You will also be contributing to exciting new research in child development and the origins of human language and communication.

What will happen to the data I provide?

The data you provide will be used alongside the data of other participants to investigate how babble contributes to brain development (and vice versa) during the first year of life. Your data will be stored securely in the University of York, Department of Language and Linguistic Science.

What about confidentiality?

Your identity will be kept strictly confidential. No real names will be used in any presentations or publications. You will be asked for your permission in writing to use excerpts from the audio or video recordings in talks or in teaching. I will not reveal your or your child's names or identity in such talks. You may take part in the study without agreeing for such use to be made of your and/or your child's recording.

Will I know the results?

If you would like to know the results, then I would be happy to send you a summary of the findings from the whole group of infants involved in the study, after the study is complete. I would also be happy to provide you with a physical/electronic copy of your completed Developmental Diary as a souvenir.

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 chair of the L&LS Ethics Committee, Márton Sóskuthy (marton.soskuthy@york.ac.uk; ***Tel:*** *(01904) 324171).*

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

Florence Oxley
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
tel: 07956 052247
email: fo507@york.ac.uk

Dr. Tamar Keren-Portnoy
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
tel: 01904 323614
email: tamar.keren-portnoy@york.ac.uk

Dr. Eytan Zweig
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD
tel: 01904 322663
email: eytan.zweig@york.ac.uk

Appendix II: Consent form

Title of project: Laterality in Babble

Lead researcher: Florence Oxley

CONSENT FORM

This form is for you to state whether or not you agree 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 you and your baby provide will be held in confidence by the research team, and your name or identifying information about you will not be mentioned in any publication? Yes ☐ No ☐

Do you understand that you may withdraw 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 your and your baby's 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 agree to take part in the study? Yes ☐ No ☐

If yes, do you agree to your interview being recorded on video? Yes ☐ No ☐

Do you agree to excerpts from your audio/video recordings and still-frames taken from the video recordings to be used in presentations or in teaching by the researcher, without disclosing your real name?

(You may take part in the study without agreeing to this). Yes ☐ No ☐

Do you agree to the researcher's keeping your contact details after the end of the current project, in order that she 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: ____

Researcher's

signature:____

Date:

____

Appendix III: Full output of mixed effects models

Model 1: modelling category

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ category + (1 | baby_name)

Data: OAI_data

REML criterion at convergence: -2234

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.6207	-0.5897	0.0014	0.6216	4.0632

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.0000753	0.008678
Residual		0.0128796	0.113488
Number of obs:	1494	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.02559	0.006697	0.3821	3.821	**0.000563** ***
Categorynb	-0.02823	0.007552	0.001488	-3.738	**0.000193** ***
Categorysm	-0.03421	0.007675	0.001489	-4.458	**0.00000889** ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Intercept)	category[nb]
category[nb]	-0.693	
category[sm]	-0.680	0.604

Model 2: modelling category and age

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ category * age_days + (1 | baby_name)

Data: OAI_data

REML criterion at convergence: -2202

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.6123	-0.6173	0.0207	0.6146	3.9650

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.0000687	0.008289
Residual		0.0127492	0.112912
Number of obs:	1494	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	-0.128	0.04361	0.001018	-2.585	**0.00986 ****
category[nb]	0.1192	0.04705	0.001381	2.533	**0.01142 ***
category[sm]	0.04813	0.04746	0.001452	1.014	0.31069
age_days	0.0004154	0.0001407	0.001195	3.208	**0.00137 ****
category[nb]:age_days	-0.0004865	0.0001552	0.001384	-3.136	**0.00175 ****
category[sm]:age_days	-0.0002456	0.0001567	0.001455	-1.568	**0.11708**

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Intercept)	category[nb]	category[sm]	age_days	category[nb]
category[nb]	-0.918				
category[sm]	-0.905	0.836			
age_days	-0.988	0.911	0.889		
category[nb]:age_days	0.893	-0.986	-0.813	-0.903	
category[sm]:age_days	0.879	-0.812	-0.986	-0.890	0.805

Model 3: modelling category and vms

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ category * vms + (1 | baby_name)

Data: OAI_data

REML criterion at convergence: -2203.6

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.7145	-0.6049	-0.0109	0.6274	4.0342

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.00013	0.0114
Residual		0.01283	0.1132
Number of obs:	1494	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.005437	0.01640	0.04894	0.332	0.7403
category[nb]	-0.01631	0.02177	0.001481	-0.749	0.4537
category[sm]	-0.0033869	0.02161	0.001483	-0.179	0.8579
vms[post]	0.03187	0.01805	0.001463	1.766	**0.0777 .**
vms[pre]	0.01074	0.01893	0.001477	0.568	0.5704
category[nb]:vms[post]	-0.01387	0.02451	0.001483	-0.566	0.5174
category[sm]: vms[post]	-0.03071	0.02442	0.001484	-1.257	0.2088
category[nb]:vms[pre]	-0.007915	0.02485	0.001481	-0.319	0.7501
category[sm]: vms[pre]	-0.03183	0.02483	0.001482	-1.282	0.2000

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Inter-cept)	category [nb]	category [sm]	vms [post]	vms [pre]	[nb] [post]	[sm] [post]	[nb] [pre]
category[nb]	-0.708							
category[sm]	-0.713	0.538						
vms[post]	-0.858	0.647	0.653					
vms[pre]	-0.809	0.609	0.614	0.729				
category[nb]:vms[post]	0.629	-0.889	-0.478	-0.728	-0.540			
category[sm]:vms[post]	0.632	-0.476	-0.886	-0.732	-0.542	0.536		
category[nb]:vms[pre]	0.618	-0.875	-0.469	-0.561	-0.756	0.777	0.415	
category[sm]: vms[pre]	0.619	-0.466	-0.869	-0.563	-0.756	0.414	0.769	0.575

Model 4: modelling category and sex

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ category * age_days + (1 | baby_name)

Data: OAI_data

REML criterion at convergence: -2213

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.6090	-0.5867	0.0011	0.6141	4.1055

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.0001033	0.01017
Residual		0.0128888	0.11353
Number of obs:	1494	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.02769	0.01072	0.1884	2.582	**0.018342**
category[nb]	-0.02751	0.01146	0.001483	-2.400	**0.016539**
category[sm]	-0.04012	0.01166	0.001486	-3.441	**0.000596**
sex[m]	0.003661	0.0141	0.2184	-0.260	0.797646
category[nb]:sex[m]	0.001168	0.01524	0.001485	-0.077	0.938935
category[sm]:sex[m]	0.01043	0.01550	0.001486	0.673	0.501067

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Intercept)	category[nb]	category[sm]	age_days	category[nb]
category[nb]	-0.648				
category[sm]	-0.635	0.596			
age_days	-0.760	0.493	0.482		
category[nb]:sex[m]	0.488	-0.752	-0.448	-0.662	
category[sm]:sex[m]	0.478	-0.449	-0.753	-0.649	0.602

Model 5: modelling utterance complexity

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ syllable_count + (1 | baby_name)

Data: OAI_data_mbp_srv

REML criterion at convergence: -528.9

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.5454	-0.5579	-0.0083	0.5462	3.4024

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.0003623	0.01903
Residual		0.0130753	0.11435
Number of obs:	369	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	1.756e-02	1.287e-02	2.492e+01	1.364	0.185
syllable_count[mono]	6.003e-04	1.722e-02	3.659e+02	0.035	0.972
syllable_count[poly]	1.661e-02	1.397e-02	3.659e+02	1.189	0.235

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

	(Intercept)	syllable_count[mono]

syllable_count[mono]	-0.528	
syllable_count[poly]	-0.66	0.486

Model 6: modelling age and utterance length

Linear mixed model fit by REML.

t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ age_days * syllable_count + (1 | baby_name)

Data: OAI_data_mbp_srv

REML criterion at convergence: -494.8

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.5698	-0.6068	-0.0374	0.6054	3.3060

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.000118	0.01086
Residual		0.012938	0.11374
Number of obs:	369	groups:	baby_name, 8

Fixed effects:

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	-1.309e-01	7.482e-02	2.823e+02	-1.750	**0.0812 .**
age_days	4.784e-04	2.388e-04	2.949e+02	2.004	**0.0460 ***
syllable_count[mono]	-9.259e-02	1.268e-01	3.626e+02	-0.730	0.4658
syllable_count[poly]	7.861e-02	9.750e-02	3.593e+02	0.806	0.4206
age_days: syllable_count[mono]	2.951e-04	4.047e-04	3.620e+02	0.729	0.4663
age_days: syllable_count[poly]	-1.934e-04	3.155e-04	3.566e+02	-0.613	0.5402

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Intercept)	age_days	syllable_count [mono]	syllable_count [poly]	age_days: syllable_count [mono]
age_days	-0.988				
syllable_count[mono]	-0.576	0.570			
syllable_count[poly]	-0.749	0.742	0.445		
age_days: syllable_count[mono]	0.570	-0.577	-0.991	-0.439	
age_days: syllable_count[poly]	0.732	-0.742	-0.435	-0.990	0.439

Model 7: modelling age and babble subtype

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']

Formula: oai ~ age_days * subtype + (1 | baby_name)

Data: OAI_data_mbp_srv

REML criterion at convergence: -494.5

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.5613	-0.6340	0.0368	0.5771	3.3896

Random effects:

Groups	Name	Variance	Std. Dev
baby_name	(Intercept)	0.0001925	0.01387
Residual		0.01290480	0.1136
Number of obs:	369	groups:	baby_name, 8

Fixed effects:

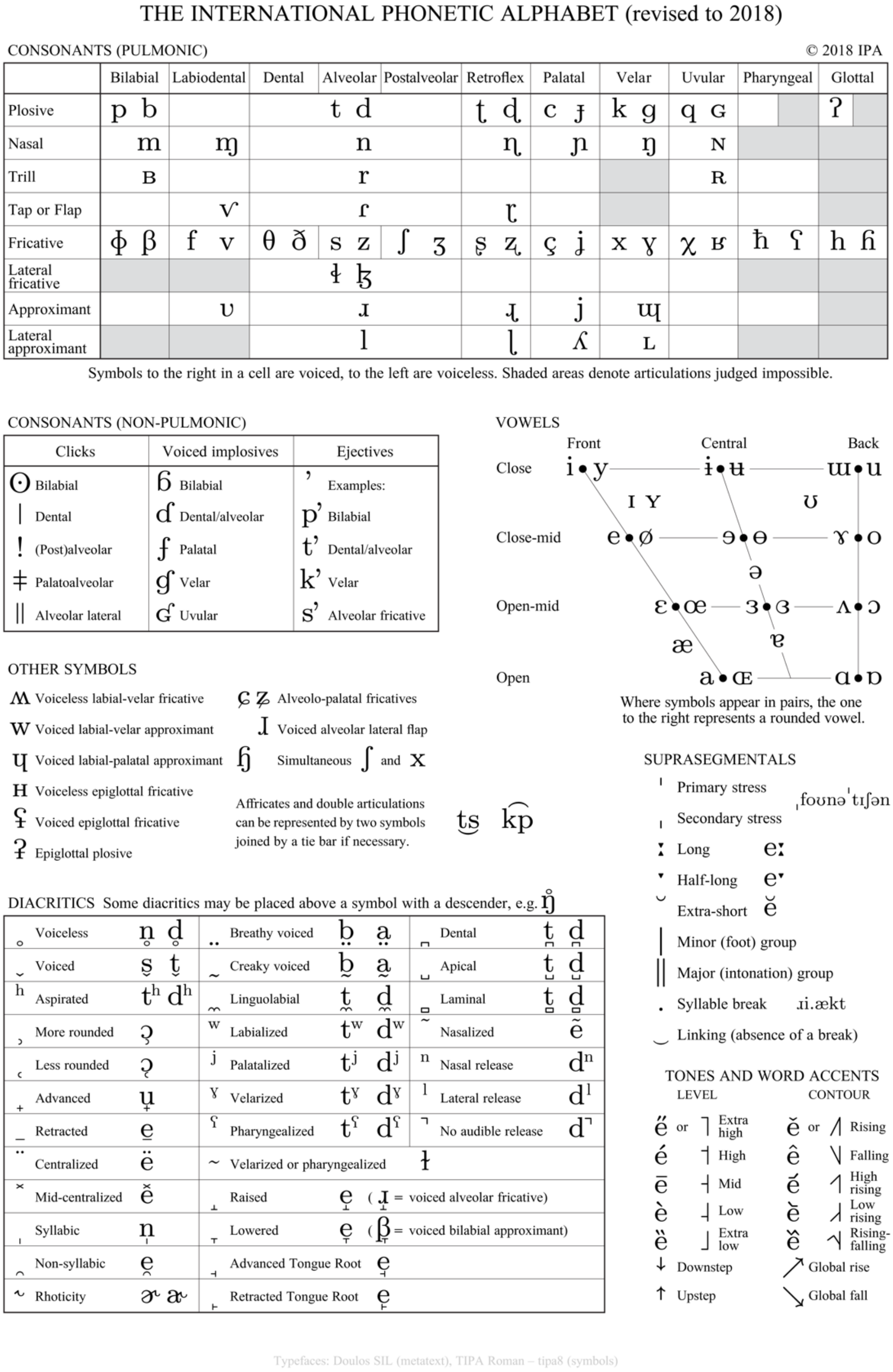
	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	-5.223e-02	9.902e-02	3.433e+02	-0.527	0.598
age_days	2.365e-04	3.230e-04	3.469e+02	0.732	0.465
subtype[singleton]	-1.620e-01	1.218e-01	3.628e+02	-1.330	0.184
subtype[variegated]	2.636e-02	1.178e-01	3.618e+02	0.224	0.823
age_days: subtype[singleton]	5.329e-04	3.944e-04	3.630e+02	1.351	0.177
age_days: subtype[variegated]	-5.826e-05	3.850e-04	3.620e+02	-0.151	0.880

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

Correlation of Fixed Effects:

	(Intercept)	age_days	subtype [singleton]	subtype [variegated]	age_days: syllable_count [singleton]
age_days	-0.990				
subtype[singleton]	-0.798	0.792			
subtype[variegated]	-0.820	0.814	0.671		
age_days:subtype[singleton]	0.799	-0.807	-0.991	-0.671	
age_days:subtype[variegated]	0.812	-0.820	-0.665	-0.991	0.676

Appendix IV: International Phonetic Alphabet



THE INTERNATIONAL PHONETIC ALPHABET (revised to 2018)

CONSONANTS (PULMONIC)

© 2018 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
ʘ Bilabial	ɓ Bilabial	ʼ Examples:
ǀ Dental	ɗ Dental/alveolar	pʼ Bilabial
ǃ (Post)alveolar	ʄ Palatal	tʼ Dental/alveolar
ǂ Palatoalveolar	ɠ Velar	kʼ Velar
ǁ Alveolar lateral	ʛ Uvular	sʼ Alveolar fricative

VOWELS

Front Central Back

Close i•y — ɨ•ʉ — ɯ•u

ɪ ʏ ʊ

Close-mid e•ø — ɘ•ɵ — ɤ•o

ə

Open-mid ɛ•œ — ɜ•ɞ — ʌ•ɔ

æ ɐ

Open a•ɶ — ɑ•ɒ

Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

ʍ Voiceless labial-velar fricative
w Voiced labial-velar approximant
ɥ Voiced labial-palatal approximant
ʜ Voiceless epiglottal fricative
ʢ Voiced epiglottal fricative
ʡ Epiglottal plosive

ɕ ʑ Alveolo-palatal fricatives
ɺ Voiced alveolar lateral flap
ɧ Simultaneous ʃ and x

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. t͡s k͡p

SUPRASEGMENTALS

ˈ Primary stress ˌfoʊnəˈtɪʃən
ˌ Secondary stress
ː Long eː
ˑ Half-long eˑ
˘ Extra-short ĕ
| Minor (foot) group
‖ Major (intonation) group
. Syllable break ɹi.ækt
‿ Linking (absence of a break)

DIACRITICS Some diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥ Voiceless	n̥ d̥	̤ Breathy voiced	b̤ a̤	̪ Dental	t̪ d̪
̬ Voiced	s̬ t̬	̰ Creaky voiced	b̰ a̰	̺ Apical	t̺ d̺
ʰ Aspirated	tʰ dʰ	̼ Linguolabial	t̼ d̼	̻ Laminal	t̻ d̻
̹ More rounded	ɔ̹	ʷ Labialized	tʷ dʷ	̃ Nasalized	ẽ
̜ Less rounded	ɔ̜	ʲ Palatalized	tʲ dʲ	ⁿ Nasal release	dⁿ
̟ Advanced	u̟	ˠ Velarized	tˠ dˠ	ˡ Lateral release	dˡ
̠ Retracted	e̠	ˤ Pharyngealized	tˤ dˤ	̚ No audible release	d̚
̈ Centralized	ë	̴ Velarized or pharyngealized	ɫ		
̽ Mid-centralized	e̽	̝ Raised	e̝ (ɹ̝ = voiced alveolar fricative)		
̩ Syllabic	n̩	̞ Lowered	e̞ (β̞ = voiced bilabial approximant)		
̯ Non-syllabic	e̯	̘ Advanced Tongue Root	e̘		
˞ Rhoticity	ɚ a˞	̙ Retracted Tongue Root	e̙		

TONES AND WORD ACCENTS

LEVEL		CONTOUR	
e̋ or ˥	Extra high	ě or ˩˥	Rising
é ˦	High	ê ˥˩	Falling
ē ˧	Mid	e᷄ ˦˥	High rising
è ˨	Low	e᷅ ˩˨	Low rising
ȅ ˩	Extra low	e᷈ ˧˦˧	Rising-falling
ꜜ	Downstep	↗	Global rise
ꜛ	Upstep	↘	Global fall

Typefaces: Doulos SIL (metatext), TIPA Roman – tipa8 (symbols)

IPA Chart, http://www.internationalphoneticassociation.org/content/ipa-chart, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2018 International Phonetic Association.

Appendix V: Transcripts

Transcript 1

The below table contains a babble transcript from Freya during one home visit made at 0;10.3 (307 days of age). Freya favoured palatals but had not yet attained a Vocal Motor Scheme at this time. Her first Vocal Motor Scheme /t, d/ emerged at 0;10.16 (320 days):

Age (days)	Babble utterance (IPA)	Supplementary notes
307	ʊɛmː	
307	ŋa	
307	nːgaga	
307	nːdɪdɪ	
307	ŋə əʊ	
307	ɪːðɪː	followed by NB shriek
307	pʰə	
307	pʰə	
307	əɹɛːjɛjɛjɛjəjɛː	
307	ɛjɛjɛ	
307	əjajajajaɹa	
307	əjajajaja	
307	ɛːːdɒɹɛ	
307	huː	

Transcript 2

The below table contains a babble transcript from Fred during one home visit made at 0;9.12 (286 days). Fred attained his first Vocal Motor Scheme /t, d/ during this home visit.

Age (days)	Babble utterance (IPA)	Supplementary notes
286	gəʊːgwəʊ	
286	gwəʔ	
286	dɪhɪ	
286	ədəːw	
286	həɹəɹaɹa	
286	had̪əd̪əʔ	
286	dəðːdɪðduðədədədəðdəd	
286	aɹəɹəɹə aɹəɹəʊ	
286	ədə	
286	adjədjçːdəʔ	
286	ə̥tə̥ də da də ədə	
286	dadədə	last syllable is squealed
286	didəd tə̥ ədədtʰə̥	
286	dəh adadɛh	
286	hɪðə	
286	də	
286	daːd ædæː	
286	dədɛdədɛdədɛdədɛdɛː	
286	dəd̪ɛd̪ɛ hadədə	growly voice
286	əðəd̪əː	
286	d̬əːː	
286	əɣwəːwˠəwˠawə	
286	ləːl hɪdə	
286	əðːdə dɪðdɛðːtʰ	
286	dədədədədədəd	some plosives fricated - saliva
286	ədədədədəd̚ dədə	some plosives fricated - saliva
286	ədədɪðːːːː	squealed towards end

Transcript 3

The below table contains a babble transcript from Orelia during one home visit made at 0;10.26 (332 days of age). Orelia was a very voluble baby and attained her first Vocal Motor Scheme /t, d/ during the first home visit after babble emergence at 0;7.2 (216 days):

Age (days)	Babble utterance (IPA)	Supplementary notes
332	ŋːə dəʊtʰ	
332	bʷe	
332	βə	
332	dəʔ	
332	bʷə	
332	tʰə̥	
332	dəʔ	
332	dadɔ	
332	daɪdə	
332	da	
332	də	
332	dɔː	
332	la	
332	də	
332	dəːdə	
332	dudəː	
332	naː	
332	dəː naː dəː	
332	əːi ndəð	
332	dəʔθ	
332	da	
332	duˑə	
332	uwəʔ	
332	wæ	
332	wəp	

332	əɹə
332	ədəɹdəʊɛʔ
332	ədədədədədədəːːː
332	ədʊɪdʊɛdʊɛːðɛː
332	əʋaʔ
332	uːɛːmoh
332	ədəː
332	βwɛː
332	dʊdɔː
332	ɪwʊ
332	uwəːː
332	dəː
332	dəː
332	uwæ
332	dəɹɛː
332	dənəuə
332	uːː uwəːː
332	əːʙəːː
332	βuːβə
332	dəʀəː
332	bwə dɔdɔː ua
332	ədədəːː
332	dəː
332	ədəʔ dədəː ɔdɛː
332	dɛːː
332	dəːː
332	dɔːdɔːdɛ
332	haːtijɔ
332	dɔiðɛdəːdəː ðɛːdɛðɛː
332	nɛʔ
332	nːɛːnɛːnɛ nɛːnɛː

332	nənɛː
332	dɛːnɛː ənɛːnɛː ənɛ
332	nɛːnaː
332	nːaːn
332	naːna
332	ba
332	aɪədə
332	læbɛba
332	da
332	la
332	əʊə
332	əβːələ
332	dɜːːð
332	dɪːð
332	idəː
332	d̪ə
332	əβə_ə
332	bebɛ
332	mə
332	bəβɛ
332	βəβɒ
332	bələləbələ
332	aba
332	da əwəːːdoʔ
332	ða
332	bɛ

Appendix VI: Vocal Motor Scheme counting

Vocal Motor Scheme count 1

The below table shows the consonant counts for Freya’s babble transcript at 0;10.3 (307 days of age) as seen in Appendix V: Transcript 1

Age	p/b	t/d	k/g	m	n	ŋ	θ/ð	h/ɦ	ɹ	j
0;10.3	2	2	1	1	2	2	1	1	3	4

Vocal Motor Scheme count 2

The below table shows the consonant counts for Fred’s babble transcript at 0;9.12 (286 days) as seen in Appendix V: Transcript 2

Age	t/d	k/g	ʔ	θ/ð	x/ɣ	h/ɦ	ɹ	l	j	w
0;9.12	20	2	3	5	1	7	2	1	1	4

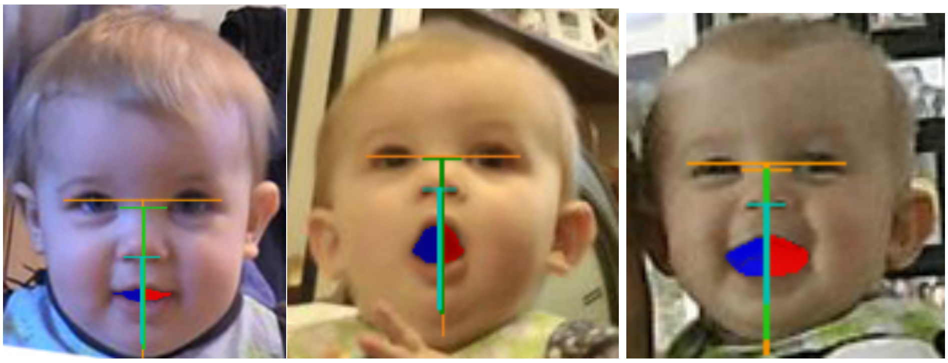
Vocal Motor Scheme count 3

The below table shows the consonant counts for Orelia’s babble transcript at 0;10.26 (332 days of age) as seen in Appendix V: Transcript 3

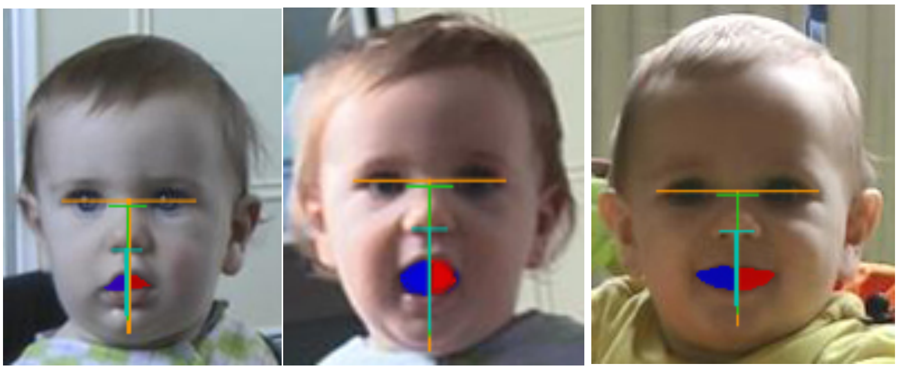
Age	p/b	t/d	ʔ	ʙ	ʀ	m	n	ŋ
0;10.26	11	44	9	1	1	2	11	1
	ɸ/β	θ/ð	h/ɦ	ʋ	ɹ	l	j	w
	7	7	2	2	2	5	1	10

Appendix VII: Still-frames

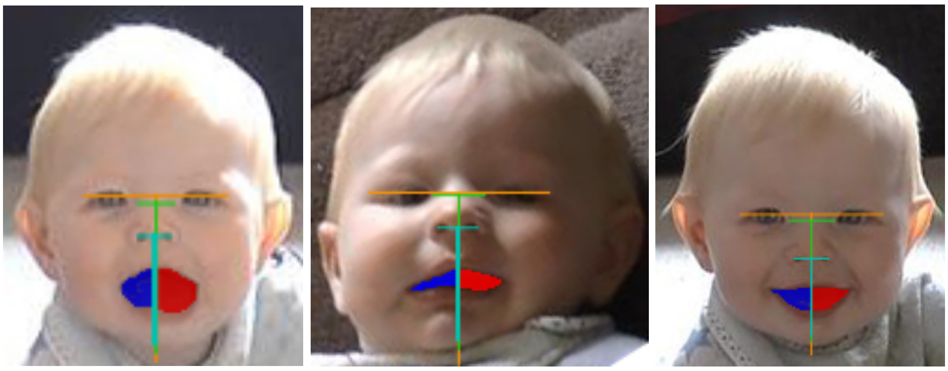
Analysed still-frames of Freya, left to right: babble, non-babble, and smile



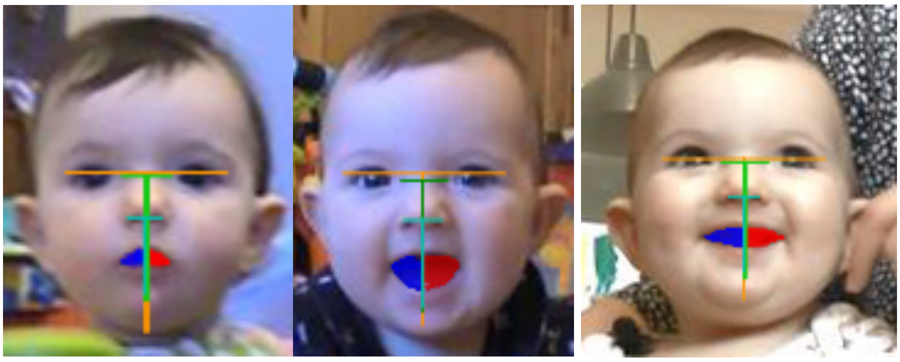
Analysed still-frames of Benji, left to right: babble, non-babble, and smile



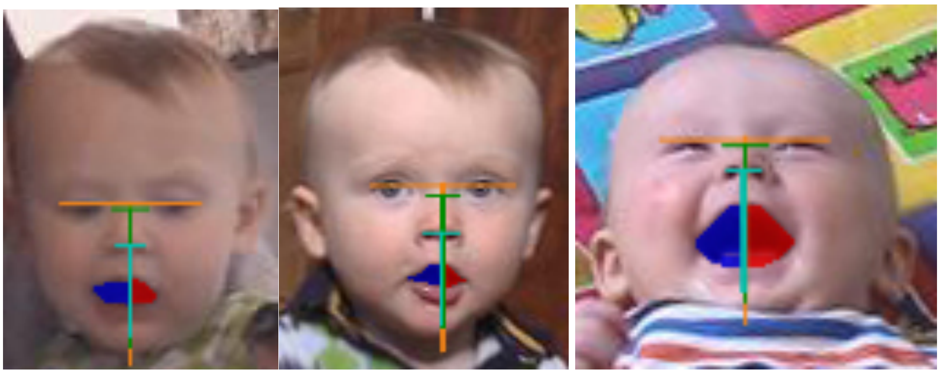
Analysed still-frames of Cameron, left to right: babble, non-babble, and smile



Analysed still-frames of Orelia, left to right: babble, non-babble, and smile

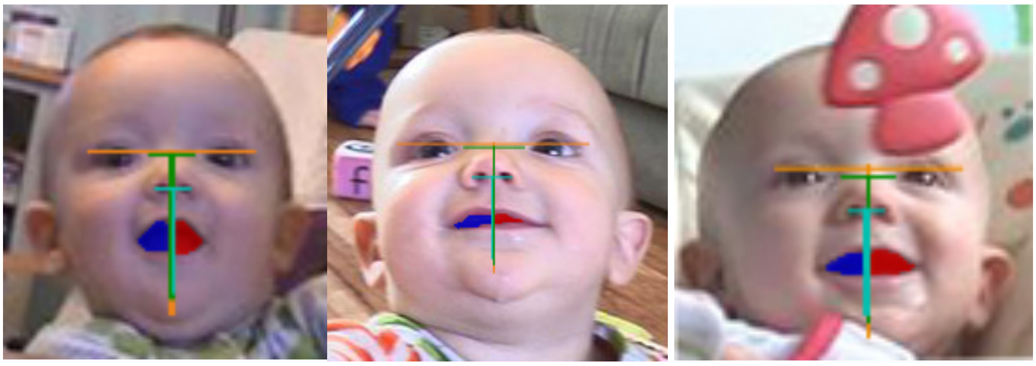


Analysed still-frames of Fred, left to right: babble, non-babble, and smile

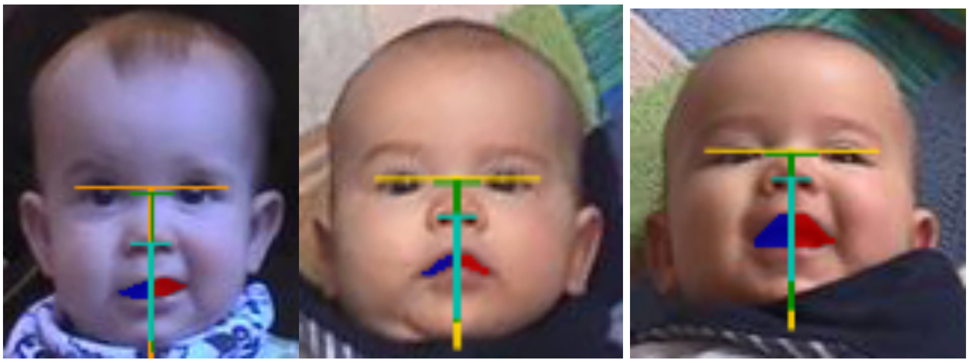


Analysed still-frames of Adelaide, left to right: babble, non-babble, and smile

Analysed still-frames of Arthur, left to right: babble, non-babble, and smile



Analysed still-frames of Leif, left to right: babble, non-babble, and smile



Appendix VIII: OAI spreadsheets

Excerpt from OAI_data

The below table illustrates the way in which the 'OAI_data' and 'OAI_data_reduced_dataset' spreadsheets are organised.

baby_name	age_days	still_name	oai	category	vms
freya	165	00.04.20.067	-0.032258065	non-babble	pre_vms
benji	182	00.19.42.420	-0.118743013	smile/laugh	pre_vms
cameron	258	00.03.47.893	0.03030303	smile/laugh	at_vms
orelia	226	00.21.23.058	0.018867925	babble	post_vms
fred	286	00.13.18.051	0.134831461	non-babble	at_vms
adelaide	336	00.01.49.931	-0.060408163	babble	post_vms
arthur	348	00.21.12.360	0.06122449	smile/laugh	post_vms
leif	258	00.22.35.063	-0.082962963	babble	pre_vms

Excerpt from OAI_data_mbp_srv

The below table illustrates the way in which the 'OAI_data_mbp_srv' spreadsheet is organised.

baby_name	age_days	still_name	oai	subtype	syllable_count
freya	194	00.02.38.685	0.04	singleton	bi
benji	252	00.24.38.942	0.10331384	variegated	poly
cameron	224	00.02.40.718	0.007294833	reduplicated	bi
orelia	251	00.24.02.725	-0.037037037	reduplicated	poly
fred	342	00.21.06.998	-0.066447609	singleton	mono
adelaide	309	00.10.12.374	0.047619048	reduplicated	poly
arthur	333	00.12.50.403	0.077922078	variegated	bi
leif	358	00.00.50.550	-0.04	singleton	mono

Appendix IX: Access to data

All data are pseudonymised. Owing to the methodology used in this study, caregiver consent has been obtained to use unfiltered images of babies' faces for the purposes of **examination only**.

A small selection of the data gathered during this study is included in the preceding Appendices for illustration and ease of reference. These include examples of analysed still-frames, example transcripts from one home visit for three different babies, an example Vocal Motor Scheme consonant count for each of these transcripts, and two excerpts from the spreadsheets used in the mixed effects models presented in in Chapter 5 and Appendix III. A larger selection of the data gathered is included as separate files accompanying the digital submission of this thesis. These data include a wider selection of still-frames before and after analysis, all transcripts of all babies' babble (used to generate Vocal Motor Scheme counts), all spreadsheets used in the mixed effects models presented in Chapter 5 and Appendix III, and all R script used to run these models and generate the visuals presented in Chapter 5 of this thesis. For reasons of space, access to the full corpus of still-frames pre- and post-analysis is available via the following link: [REDACTED FOR PARTICIPANT PRIVACY IN LINE WITH ETHICS APPROVAL]. A .txt file entitled 'Oxley_PhD submission_README.txt' is also supplied alongside the digital submission of this thesis, which explains how the accompanying data files relate to one another and to the thesis.

This data is to be used for the purposes of thesis examination only.

For reasons of ethics, this data must not be used for any other purpose and must not be copied, edited, reproduced, or shared with any party besides the researcher, the supervisors, and the examiners.

Many thanks for your co-operation.

Florence Oxley

fo507@york.ac.uk

References

Abbs, J. H., & DePaul, R. (1998). Motor cortex fields and speech movements: Simple dual control is implausible. *Behavioral and Brain Sciences*, *21*(4), 511-512.

Alcock, K. J. (2006). The development of oral motor control and language. *Down syndrome research and practice*, *11*(1), 1-8.

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*(18), 4909-4917. https://doi.org/10.1073/pnas.1603205113

Andrew, R. (1998). Cyclicity in speech derived from call repetition rather than from intrinsic cyclicity of ingestion. *Behavioral and Brain Sciences*, *21*(4), 513-514.

Annett, M. (1996). In Defence of the Right Shift Theory. *Perceptual and Motor Skills*, *82*(1), 115-137. https://doi.org/10.2466/pms.1996.82.1.115

Arcadi, A. C. (2003). Is gestural communication more sophisticated than vocal communication in wild chimpanzees? *Behavioral and Brain Sciences*, *26*(2), 210-211.

Arnold, K., & Zuberbühler, K. (2006). Semantic combinations in primate calls. *Nature*, *441*(7091), 303-303.

babble, n. In. (n.d.). *Oxford English Dictionary*. Retrieved 28th June 2022, from https://www.oed.com/view/Entry/14198?rskey=PWmunS&result=1

Barton, K. (2022). *MuMIn: Multi-Model Inference. R package version 1.46.0.* https://CRAN.R-project.org/package=MuMIn

Bastiaansen, J. A. C. J., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2391-2404. https://doi.org/10.1098/rstb.2009.0058

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. . *Journal of Statistical Software 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bates, E. (1999). Plasticity, localization and language development. In S. H. Broman & J. M. Fletcher (Eds.), *The changing nervous system: Neurobehavioral consequences of early brain disorders* (pp. 214-253). Oxford Univesity Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The Emergence of Symbols: Cognition and communication in infancy.* Academic Press.

Bates, E. A. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language*, *88*(2), 248-253.

Beaton, A. A. (2003). Going for Broca? I wouldn't bet on it! *Behavioral and Brain Sciences*, *26*(2), 212-213.

Beaudet, A. (2017). The emergence of language in the hominin lineage: perspectives from fossil endocasts. *Frontiers in Human Neuroscience*, *11*, 427.

Becker, Y., Phelipon, R., Sein, J., Velly, L., Renaud, L., & Meguerditchian, A. (2022). Planum temporale grey matter volume asymmetries in newborn monkeys (Papio anubis). *Brain Structure and Function*, *227*(2), 463-468.

Becker, Y., Sein, J., Velly, L., Giacomino, L., Renaud, L., Lacoste, R., Anton, J.-L., Nazarian, B., Berne, C., & Meguerditchian, A. (2021). Early Left-Planum Temporale Asymmetry in newborn monkeys (Papio anubis): A longitudinal structural MRI study at two stages of development. *NeuroImage*, *227*, 117575.

Bermejo-Fenoll, A., Panchón-Ruíz, A., & Sánchez del Campo, F. (2021). Hypothesis about the appearance of the vermillion border of the lips in Homo sapiens. *Translational Research in Anatomy*, *24*, 100106. https://doi.org/https://doi.org/10.1016/j.tria.2020.100106

Best, C. T., & Queen, H. F. (1989). Baby, it's in your smile: Right hemiface bias in infant emotional expressions. *Developmental Psychology*, *25*(2), 264-276. https://doi.org/10.1037/0012-1649.25.2.264

Bigand, F., Prigent, E., Berret, B., & Braffort, A. (2021). How fast is Sign Language? A reevaluation of the kinematic bandwidth using motion capture. 2021 29th European Signal Processing Conference (EUSIPCO),

Binder, J. R. (2017). Current controversies on Wernicke's area and its role in language. *Current neurology and neuroscience reports*, *17*(8), 1-10.

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*(4), 1239-1247.

Bishop, D. V. M. (2013). Cerebral Asymmetry and Language Development: Cause, Correlate, or Consequence? *Science*, *340*(6138), 1230531. https://doi.org/doi:10.1126/science.1230531

Blaser, S., Propst, E. J., Martin, D., Feigenbaum, A., James, A. L., Shannon, P., & Papsin, B. C. (2006). Inner Ear Dysplasia is Common in Children With Down Syndrome (trisomy 21). *The Laryngoscope*, *116*(12), 2113-2119. https://doi.org/10.1097/01.mlg.0000245034.77640.4f

Bleile, K. M., Stark, R. E., & McGowan, J. S. (1993). Speech development in a child after decannulation: Further evidence that babbling facilitates later speech development. *Clinical Linguistics & Phonetics*, *7*(4), 319-337.

Bloom, K. (1975). Social elicitation of infant vocal behavior. *Journal of Experimental Child Psychology*, *20*(1), 51-58. https://doi.org/https://doi.org/10.1016/0022-0965(75)90025-9

Bloom, K. (1977). Patterning of infant vocal behavior. *Journal of Experimental Child Psychology*, *23*(3), 367-377. https://doi.org/https://doi.org/10.1016/0022-0965(77)90032-7

Bloom, K. (1998). The missing link's missing link: Syllabic vocalizations at 3 months of age. *Behavioral and Brain Sciences*, *21*(4), 514-515.

Bludau, S., Eickhoff, S. B., Mohlberg, H., Caspers, S., Laird, A. R., Fox, P. T., Schleicher, A., Zilles, K., & Amunts, K. (2014). Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage*, *93*, 260-275. https://doi.org/https://doi.org/10.1016/j.neuroimage.2013.05.052

Bogolepova, N., & Malofeeva, L. I. (2001). Characteristics of the development of speech motor areas 44 and 45 in the left and right hemispheres of the human brain in early post-natal ontogenesis. *Neurosci Behav Physiol*, *31*(4), 349-354. https://doi.org/10.1023/a:1010468007795

Bohm, L. A., Nelson, M. E., Driver, L. E., & Green, G. E. (2010). Babbling, vegetative function, and language development after cricotracheal resection in aphonic children. *The Laryngoscope*, *120*(12), 2494-2497.

Borod, J. C., Kent, J., Koff, E., Martin, C., & Alpert, M. (1988). Facial asymmetry while posing positive and negative emotions: Support for the right hemisphere hypothesis. *Neuropsychologia*, *26*(5), 759-764. https://doi.org/https://doi.org/10.1016/0028-3932(88)90013-9

Borod, J. C., Koff, E., & White, B. (1983). Facial asymmetry in posed and spontaneous expressions of emotion. *Brain and Cognition*, *2*(2), 165-175. https://doi.org/https://doi.org/10.1016/0278-2626(83)90006-4

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. J., & Frith, D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language A positron emission tomography activation study. *Brain*, *117*(6), 1241-1253. https://doi.org/10.1093/brain/117.6.1241

Bradshaw, J. L. (2003). Gesture in language evolution: Could I but raise my hand to it! *Behavioral and Brain Sciences*, *26*(2), 213-214.

Broadfield, D. C., Holloway, R. L., Mowbray, K., Silvers, A., Yuan, M. S., & Márquez, S. (2001). Endocast of Sambungmacan 3 (Sm 3): A newHomo erectusfrom Indonesia. *The Anatomical Record*, *262*(4), 369-379. https://doi.org/10.1002/ar.1047

Brunet, M., Guy, F., Pilbeam, D., Mackaye, H. T., Likius, A., Ahounta, D., Beauvilain, A., Blondel, C., Bocherens, H., Boisserie, J.-R., De Bonis, L., Coppens, Y., Dejax, J., Denys, C., Duringer, P., Eisenmann, V., Fanone, G., Fronty, P., Geraads, D., . . . Zollikofer, C. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature*, *418*(6894), 145-151. https://doi.org/10.1038/nature00879

Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, *25*(5), 663-678. https://doi.org/10.1207/s15516709cog2505_2

Bukowska, M., Essick, G.K. & Trulsson, M. Functional properties of low-threshold mechanoreceptive afferents in the human labial mucosa. *Exp Brain Res* **201**, 59–64 (2010). https://doi.org/10.1007/s00221-009-2005-0

Bunton, K. (2008). Speech versus nonspeech: different tasks, different neural organization. *Semin Speech Lang*, *29*(4), 267-275. https://doi.org/10.1055/s-0028-1103390

Cantalupo, C., & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes: A region of the ape brain is uncannily similar to one linked with speech in humans. *Nature*, *414*(6863), 505-505. https://doi.org/10.1038/35107134

Cantiani, C., Ortiz-Mantilla, S., Riva, V., Piazza, C., Bettoni, R., Musacchia, G., Molteni, M., Marino, C., & Benasich, A. A. (2019). Reduced left-lateralized pattern of event-related EEG oscillations in infants at familial risk for language and learning impairment.

NeuroImage: Clinical, *22*, 101778. https://doi.org/https://doi.org/10.1016/j.nicl.2019.101778

Carmeli, E. (2015). Anxiety in the Elderly Can be a Vestibular Problem. *Frontiers in Public Health*, 3. https://doi.org/10.3389/fpubh.2015.00216

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience*, *13*(11), 1428-1432. https://doi.org/10.1038/nn.2641

Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., Raz, A., Kaipio, M.-L., Hiltunen, J., & Aaltonen, O. (2002). Speech sounds learned by sleeping newborns. *Nature*, *415*(6872), 599-600.

Chi, J. G., Dooling, E. C., & Gilles, F. H. (1977). Left-Right Asymmetries of the Temporal Speech Areas of the Human Fetus. *Archives of Neurology*, *34*(6), 346-348. https://doi.org/10.1001/archneur.1977.00500180040008

Chomsky, N. & Halle, M. (1968). The Sound Pattern of English. New York: Harper & Row.

Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, J. G., Mostofsky, S. H., Pekar, J. J., Denckla, M. B., & Cutting, L. E. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language*, *98*(2), 150-158. https://doi.org/https://doi.org/10.1016/j.bandl.2006.04.007

Cochet, H., & Vauclair, J. (2012). Hand preferences in human adults: Non-communicative actions versus communicative gestures. *Cortex*, *48*(8), 1017-1026. https://doi.org/10.1016/j.cortex.2011.03.016

Cook, N. D. (2003). Hemispheric dominance has its origins in the control of the midline organs of speech. *Behavioral and Brain Sciences*, *26*(2), 216-217.

Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav Brain Sci*, *26*(2), 199-208; discussion 208-160. https://doi.org/10.1017/s0140525x03000062

Corballis, M. C. (2008). Of mice and men - and lopsided birds. *Cortex*, *44*(1), 3-7. https://doi.org/10.1016/j.cortex.2007.10.001

Corballis, M. C. (2009). The evolution and genetics of cerebral asymmetry. *Philos Trans R Soc Lond B Biol Sci*, *364*(1519), 867-879. https://doi.org/10.1098/rstb.2008.0232

Corballis, M. C. (2013). Early signs of brain asymmetry. *Trends Cogn Sci*, *17*(11), 554-555. https://doi.org/10.1016/j.tics.2013.09.008

Corballis, M. C. (2014b). Left Brain, Right Brain: Facts and Fantasies. *PLoS Biology*, *12*(1), e1001767. https://doi.org/10.1371/journal.pbio.1001767

Corballis, M. C. (2015). What's left in language? Beyond the classical model. *Ann N Y Acad Sci*, *1359*, 14-29. https://doi.org/10.1111/nyas.12761

Corballis, M. C. (2017). The Evolution of Lateralized Brain Circuits. *Front Psychol*, *8*, 1021. https://doi.org/10.3389/fpsyg.2017.01021

Corbetta, D. (2003). Right-handedness may have come first: Evidence from studies in human infants and nonhuman primates. *Behavioral and Brain Sciences*, *26*(2), 217-218.

Crespi, B., Read, S., & Hurd, P. (2017). Segregating polymorphisms of FOXP2 are associated with measures of inner speech, speech fluency and strength of handedness in a healthy population. *Brain and Language*, *173*, 33-40. https://doi.org/https://doi.org/10.1016/j.bandl.2017.06.002

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, *110*(3), 221-243.

Crow, T. J. (2008). The 'big bang' theory of the origin of psychosis and the faculty of language. *Schizophr Res*, *102*(1-3), 31-52. https://doi.org/10.1016/j.schres.2008.03.010

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *NeuroImage*, *54*(4), 2808-2821. https://doi.org/https://doi.org/10.1016/j.neuroimage.2010.10.069

Davis, B. L., & Macneilage, P. F. (1994). Organization of babbling: A case study. *Language and speech*, *37*(4), 341-355.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *J Speech Hear Res*, *38*(6), 1199-1211. https://doi.org/10.1044/jshr.3806.1199

de Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 353-363). Springer.

de Boysson-Bardies, B., Bacri, N., Sagart, L., & Poizat, M. (1981). Timing in late babbling. *Journal of Child Language*, *8*(3), 525-539.

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

de Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, *67*(2), 297-319.

De Carli, D., Garreffa, G., Colonnese, C., Giulietti, G., Labruna, L., Briselli, E., Ken, S., Macrì, M. A., & Maraviglia, B. (2007). Identification of activated regions during a language task. *Magnetic Resonance Imaging*, *25*(6), 933-938. https://doi.org/https://doi.org/10.1016/j.mri.2007.03.031

DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, *17*(2), 159-164. https://doi.org/https://doi.org/10.1016/0163-6383(94)90051-5

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebral Cortex*, *25*(11), 4596-4609.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600), 2013-2015. https://doi.org/10.1126/science.1077066

Dehaene-Lambertz, G., Montavont, A., Jobert, A., Allirol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language*, *114*(2), 53-65. https://doi.org/https://doi.org/10.1016/j.bandl.2009.09.003

Deignan, A., Candarli, D., & Oxley, F. A. R. (2022). *The Linguistic Challenge of the Transition to Secondary School: A Corpus Study of Academic Language* (1st ed.). Routledge. https://doi.org/10.4324/9781003081890

Delson, E., Harvati, K., Reddy, D., Marcus, L. F., Mowbray, K., Sawyer, G. J., Jacob, T., & Márquez, S. (2001). The Sambungmacan 3 Homo erectus calvaria: A comparative morphometric and morphological analysis. *The Anatomical Record*, *262*(4), 380-397. https://doi.org/10.1002/ar.1048

Demuru, E., Ferrari, P. F., & Palagi, E. (2018). Is birth attendance a uniquely human feature? New evidence suggests that Bonobo females protect and support the parturient. *Evolution and Human Behavior*, *39*(5), 502-510. https://doi.org/https://doi.org/10.1016/j.evolhumbehav.2018.05.003

DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development*, *34*(4), 590-601.

DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development*, *36*(4), 642-649.

Deppe, M., Knecht, S., Papke, K., Lohmann, H., Fleischer, H., Heindel, W., Ringelstein, E. B., & Henningsen, H. (2000). Assessment of hemispheric language lateralization: a comparison between fMRI and fTCD. *J Cereb Blood Flow Metab*, *20*(2), 263-268. https://doi.org/10.1097/00004647-200002000-00006

Desmurget, M., Richard, N., Harquel, S., Baraduc, P., Szathmari, A., Mottolese, C., & Sirigu, A. (2014). Neural representations of ethologically relevant hand/mouth synergies in the human precentral gyrus. *Proceedings of the National Academy of Sciences*, *111*(15), 5718-5722.

Doi, H., Nishitani, S., & Shinohara, K. (2013). NIRS as a tool for assaying emotional function in the prefrontal cortex. *Frontiers in Human Neuroscience*, *7*, 770.

Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*(6605), 159-161. https://doi.org/10.1038/384159a0

Dum, R. P., & Strick, P. L. (2003). An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *J Neurophysiol*, *89*(1), 634-639. https://doi.org/10.1152/jn.00626.2002

Dusseldorp, G. L., & Lombard, M. (2021). Constraining the Likely Technological Niches of Late Middle Pleistocene Hominins with Homo naledi as Case Study. *Journal of Archaeological Method and Theory*, *28*(1), 11-52. https://doi.org/10.1007/s10816-020-09501-7

Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation, 19,* 207–283.

Ekman, P. (1992). Facial expressions of emotion: an old controversy and new findings. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *335*(1273), 63-69.

Ejiri, K. (1998). Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica*, *55*(4), 226-237.

Ejiri, K., & Masataka, N. (2001). Co-occurences of preverbal vocal behavior and motor action in early infancy. *Developmental Science*, *4*(1), 40-48. https://doi.org/10.1111/1467-7687.00147

Elbers, L. (1982). Operating principles in repetitive babbling: a cognitive continuity approach. *Cognition*, *12*(1), 45-63. https://doi.org/10.1016/0010-0277(82)90029-4

Elbers, L. (2000). An output-as-input hypothesis in language acquisition. In P. Broeder & J. Murre (Eds.), *Models of Language Acquisition: Inductive and deductive approaches* (pp. 244-271). Oxford University Press.

Elbers, L., & Ton, J. (1985). Play pen monologues: The interplay of words and babbles in the first words period. *Journal of Child Language*, *12*(3), 551-565.

Elbers, L., & Wijnen, F. (1992). Effort, Production Skill, and Language Learning. In C. A. Fergusion, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 337-368). York Press.

Elmer, S., Hänggi, J., & Jäncke, L. (2016). Interhemispheric transcallosal connectivity between the left and right planum temporale predicts musicianship, performance in temporal speech processing, and functional specialization. *Brain Structure and Function*, *221*(1), 331-344.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998). 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends in cognitive sciences*, *2*(1), 31-37.

Esling, J. H. (2012). The Articulatory Function of the Larynx and the Origins of Speech. *Annual Meeting of the Berkeley Linguistics Society*, *38*, 121. https://doi.org/10.3765/bls.v38i0.3325

Esseily, R., Jacquet, A. Y., & Fagard, J. (2011). Handedness for grasping objects and pointing and the development of language in 14-month-old infants. *Laterality*, *16*(5), 565-585. https://doi.org/10.1080/1357650x.2010.499911

Fadiga, L., & Craighero, L. (2006). Hand actions and speech representation in Broca's area. *Cortex*, *42*(4), 486-490. https://doi.org/10.1016/s0010-9452(08)70383-6

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608-2611. https://doi.org/10.1152/jn.1995.73.6.2608

Fagan, M. K. (2009). Mean Length of Utterance before words and grammar: Longitudinal trends and developmental implications of infant vocalizations. *Journal of Child Language*, *36*(3), 495-527. https://doi.org/10.1017/s0305000908009070

Fagan, M. K., & Iverson, J. M. (2007). The Influence of Mouthing on Infant Vocalization. *Infancy*, *11*(2), 191-202. https://doi.org/10.1111/j.1532-7078.2007.tb00222.x

Fedurek, P., Slocombe, K. E., Hartel, J. A., & Zuberbuhler, K. (2015a). Chimpanzee lip-smacking facilitates cooperative behaviour. *Sci Rep*, *5*, 13460. https://doi.org/10.1038/srep13460

Fedurek, P., Slocombe, K. E., & Zuberbühler, K. (2015b). Chimpanzees communicate to two different audiences during aggressive interactions. *Animal Behaviour*, *110*, 21-28. https://doi.org/https://doi.org/10.1016/j.anbehav.2015.09.010

Ferguson, C. A., & Farwell, C. B. (1975). Words and Sounds in Early Language Acquisition. *Language*, *51*(2), 419-439.

Fernández-Carriba, S., Loeches, A., Morcillo, A., & Hopkins, W. D. (2002a). Functional asymmetry of emotions in primates: new findings in chimpanzees. *Brain Research Bulletin*, *57*(3), 561-564. https://doi.org/https://doi.org/10.1016/S0361-9230(01)00685-2

Fernández-Carriba, S., Loeches, A., Morcillo, A., & Hopkins, W. D. (2002b). Asymmetry in facial expression of emotions by chimpanzees. *Neuropsychologia*, *40*(9), 1523-1533. https://doi.org/https://doi.org/10.1016/S0028-3932(02)00028-3

Fernández, G., Weis, S., Stoffel-Wagner, B., Tendolkar, I., Reuber, M., Beyenburg, S., Klaver, P., Fell, J., De Greiff, A., Ruhlmann, J., Reul, J., & Elger, C. E. (2003). Menstrual Cycle-Dependent Neural Plasticity in the Adult Human Brain Is Hormone, Task, and Region Specific. *The Journal of Neuroscience*, *23*(9), 3790-3795. https://doi.org/10.1523/jneurosci.23-09-03790.2003

Ferre, C. L., Babik, I., & Michel, G. F. (2010). Development of infant prehension handedness: A longitudinal analysis during the 6- to 14-month age period. *Infant Behavior and Development*, *33*(4), 492-502. https://doi.org/https://doi.org/10.1016/j.infbeh.2010.06.002

Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F., & Hammerschmidt, K. (2020). Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proceedings of the Royal Society B*, *287*(1941), 20202531.

Fischer, J., Wheeler, B. C., & Higham, J. P. (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Current Biology*, *25*(21), R1028-R1029.

Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., & Tsuji, S. (2018). Symbouki: a meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science*, *21*(5), e12659. https://doi.org/10.1111/desc.12659

Fox, N. A., & Davidson, R. J. (1986). Taste-elicited changes in facial signs of emotion and the asymmetry of brain electrical activity in human newborns. *Neuropsychologia*, *24*(3), 417-42. https://doi.org/10.1016/0028-3932(86)90028-X

Fox, N. A., & Davidson, R. J. (1988). Patterns of brain electrical activity during facial signs of emotion in 10-month-old infants. *Developmental Psychology*, *24*(2), 230-236. https://doi.org/10.1037/0012-1649.24.2.230

Frayer, D. W., Clarke, R. J., Fiore, I., Blumenschine, R. J., Pérez-Pérez, A., Martinez, L. M., Estebaranz, F., Holloway, R., & Bondioli, L. (2016). OH-65: The earliest evidence for right-handedness in the fossil record. *Journal of human evolution*, *100*, 65-72. https://doi.org/https://doi.org/10.1016/j.jhevol.2016.07.002

Friend, M. (2004). On the Implications of Curvilinear Trajectories for Cognitive Development. *Journal of Cognition and Development*, *5*(1), 103-108. https://doi.org/10.1207/s15327647jcd0501_8

Frost, J. A., Binder, J. R., Springer, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. *Brain*, *122*(2), 199-208. https://doi.org/10.1093/brain/122.2.199

Futagi, Y., Yanagihara, K., Mogami, Y., Ikeda, T., & Suzuki, Y. (2013). The babkin reflex in infants: clinical significance and neural mechanism. *Pediatr Neurol*, *49*(3), 149-155. https://doi.org/10.1016/j.pediatrneurol.2013.04.005

Gannon, P. J., Holloway, R. L., Broadfield, D. C., & Braun, A. R. (1998). Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science*, *279*(5348), 220-222. https://doi.org/10.1126/science.279.5348.220

Garnier, S., Ross, N., Rudis, R., Camargo, A. P., Sciaini, M., & Scherer, C. (2021). *Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.2.* https://sjmgarnier.github.io/viridis/

Gazzaniga, M. S., & Smylie, C. S. (1983). Facial recognition and brain asymmetries: clues to underlying mechanisms. *Ann Neurol*, *13*(5), 536-540. https://doi.org/10.1002/ana.410130511

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gentilucci, M., Campione, G. C., De Stefani, E., & Innocenti, A. (2012). Is the coupled control of hand and mouth postures precursor of reciprocal relations between gestures and words? *Behavioural Brain Research*, *233*(1), 130-140.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*(7), 949-960.

Gershkoff-Stowe, L., & Thelen, E. (2004). U-Shaped Changes in Behavior: A Dynamic Systems Perspective. *Journal of Cognition and Development*, *5*(1), 11-36. https://doi.org/10.1207/s15327647jcd0501_2

Ghazanfar, A. A., & Katz, D. B. (1998). Distributed neural substrates and the evolution of speech production. *Behavioral and Brain Sciences*, *21*(4), 516-517.

Ghazanfar, A. A., & Takahashi, D. Y. (2014). Facial expressions and the evolution of the speech rhythm. *Journal of cognitive neuroscience*, *26*(6), 1196-1207.

Glanville, B. B., Best, C. T., & Levenson, R. (1977). A cardiac measure of cerebral asymmetries in infant auditory perception. *Developmental Psychology*, *13*(1), 54-59. https://doi.org/10.1037/0012-1649.13.1.54

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, *22*(2), 133-139.

Goldin-Meadow, S. (2004). U-Shaped Changes Are in the Eye of the Beholder. *Journal of Cognition and Development*, *5*(1), 109-111. https://doi.org/10.1207/s15327647jcd0501_9

Graven, S. N., & Browne, J. V. (2008). Auditory Development in the Fetus and Infant. *Newborn and Infant Nursing Reviews*, *8*(4), 187-193. https://doi.org/https://doi.org/10.1053/j.nainr.2008.10.010

Graves, R., Goodglass, H., & Landis, T. (1982). Mouth asymmetry during spontaneous speech. *Neuropsychologia*, *20*(4), 371-381. https://doi.org/https://doi.org/10.1016/0028-3932(82)90037-9

Graves, R., & Landis, T. (1990). Asymmetry in Mouth Opening During Different Speech Tasks. *International Journal of Psychology*, *25*(2), 179-189. https://doi.org/10.1080/00207599008247856

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348-353. https://doi.org/https://doi.org/10.1016/S0166-2236(02)02191-4

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*(6), 509-516.

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, *19*(4), 385-408.

Grotheer, M., Rosenke, M., Wu, H., Kular, H., Querdasi, F. R., Natu, V. S., Yeatman, J. D., & Grill-Spector, K. (2022). White matter myelination during early infancy is linked to spatial gradients and myelin content at birth. *Nature communications*, *13*(1), 1-12.

Häberling, I. S., & Corballis, M. C. (2016). Cerebellar asymmetry, cortical asymmetry and handedness: Two independent networks. *Laterality*, *21*(4-6), 397-414. https://doi.org/10.1080/1357650x.2015.1110161

Häberling, I. S., Corballis, P. M., & Corballis, M. C. (2016). Language, gesture, and handedness: Evidence for independent lateralized networks. *Cortex*, *82*, 72-85. https://doi.org/10.1016/j.cortex.2016.06.003

Hahn, W. K. (1987). Cerebral lateralization of function: From infancy through childhood. *Psychological Bulletin*, *101*(3), 376-392. https://doi.org/10.1037/0033-2909.101.3.376

Harpaz, Y., Levkovitz, Y., & Lavidor, M. (2009). Lexical ambiguity resolution in Wernicke's area and its right homologue. *Cortex*, *45*(9), 1097-1103.

Hauser, M. D., & Andersson, K. (1994). Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proceedings of the National Academy of Sciences*, *91*(9), 3946-3948. https://doi.org/10.1073/pnas.91.9.3946

Hausmann, M., Behrendt-Körbitz, S., Kautz, H., Lamm, C., Radelt, F., & Güntürkün, O. (1998). Sex differences in oral asymmetries during word repetition. *Neuropsychologia*, *36*(12), 1397-1402. https://doi.org/10.1016/s0028-3932(98)00027-x

Hayakawa, K., Konishi, Y., Kuriyama, M., Konishi, K., & Matsuda, T. (1991). Normal brain maturation in MRI. *European Journal of Radiology*, *12*(3), 208-215. https://doi.org/10.1016/0720-048x(91)90074-6

Hecht, E. E., & Parr, L. A. (2015). The chimpanzee mirror system and the evolution of frontoparietal circuits for action observation and social learning.

Hering-Hanit, R., Achiron, R., Lipitz, S., & Achiron, A. (2001). Asymmetry of fetal cerebral hemispheres: in utero ultrasound study. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *85*(3), 194F-196. https://doi.org/10.1136/fn.85.3.f194

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402. https://doi.org/10.1038/nrn2113

Hodgson, J. C., Hirst, R. J., & Hudson, J. M. (2016). Hemispheric speech lateralisation in the developing brain is related to motor praxis ability. *Dev Cogn Neurosci*, *22*, 9-17. https://doi.org/10.1016/j.dcn.2016.09.005

Hodgson, J. C., Richardson, D., & Hudson, J. M. (2021). The relationship between lateralization patterns from sequence based motor tasks and hemispheric speech dominance. *Neuropsychology*, *35*(2), 157-171. https://doi.org/10.1037/neu0000702

Hoeksema, N., Verga, L., Mengede, J., Van Roessel, C., Villanueva, S., Salazar-Casals, A., Rubio-Garcia, A., Ćurčić-Blake, B., Vernes, S. C., & Ravignani, A. (2021). Neuroanatomy of the grey seal brain: bringing pinnipeds into the neurobiological study of vocal learning. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20200252.

Holloway, R. L. (1983). Human paleontological evidence relevant to language behavior. *Human neurobiology*, *2*(3), 105-114.

Holloway, R. L. (2003). Was a manual gesturing stage really necessary? *Behavioral and Brain Sciences*, *26*(2), 223-224.

Holloway, R. L., Hurst, S. D., Garvin, H. M., Schoenemann, P. T., Vanti, W. B., Berger, L. R., & Hawks, J. (2018). Endocast morphology of Homo naledi from the Dinaledi chamber, South Africa. *Proceedings of the National Academy of Sciences*, *115*(22), 5738-5743.

Holowka, S., & Petitto, L. A. (2002). Left hemisphere cerebral specialization for babies while babbling. *Science*, *297*(5586), 1515. https://doi.org/10.1126/science.1074941

Holowka, S., & Petitto, L. A. (2002a). Left Hemisphere Cerebral Specialization for Babies While Babbling. *Science*, *297*(5586), 1515-1515. https://doi.org/doi:10.1126/science.1074941

Hook-Costigan, M. A., & Rogers, L. J. (1998). Lateralized use of the mouth in production of vocalizations by marmosets. *Neuropsychologia*, *36*(12), 1265-1273. https://doi.org/10.1016/s0028-3932(98)00037-2

Hopkins, W. D., & Cantalupo, C. (2003). Brodmann's area 44, gestural communication, and the emergence of right handedness in chimpanzees. *Behavioral and Brain Sciences*, *26*(2), 224-225.

Hopkins, W. D., & Cantero, M. (2003). From hand to mouth in the evolution of language: The influence of vocal behavior on lateralized hand use in manual gestures by chimpanzees (Pan troglodytes). *Developmental Science*, *6*(1), 55-61.

Hori, M. (1993). Frequency-Dependent Natural Selection in the Handedness of Scale-Eating Cichlid Fish. *Science*, *260*(5105), 216-219. https://doi.org/doi:10.1126/science.260.5105.216

Hostetter, A. B., Cantero, M., & Hopkins, W. D. (2001). Differential use of vocal and gestural communication by chimpanzees (Pan troglodytes) in response to the attentional status of a human (Homo sapiens). *J Comp Psychol*, *115*(4), 337-343. https://doi.org/10.1037//0735-7036.115.4.337

Intrapiromkul, J., Aygun, N., Tunkel, D. E., Carone, M., & Yousem, D. M. (2012). Inner ear anomalies seen on CT images in people with Down syndrome. *Pediatric Radiology*, *42*(12), 1449-1455. https://doi.org/10.1007/s00247-012-2490-3

Inubushi, T., & Sakai, K. L. (2013). Functional and anatomical correlates of word-, sentence-, and discourse-level integration in sign language. *Frontiers in Human Neuroscience*, *7*, 681.

Iverson, J. M., & Fagan, M. K. (2004). Infant Vocal-Motor Coordination: Precursor to the Gesture-Speech System? *Child Development*, *75*(4), 1053-1066. https://doi.org/10.1111/j.1467-8624.2004.00725.x

Iverson, J. M., Hall, A. J., Nickel, L., & Wozniak, R. H. (2007). The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and Language*, *101*(3), 198-207. https://doi.org/10.1016/j.bandl.2006.11.004

Iverson, J. M., & Thelen, E. (1999). Hand, mouth and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies*, *6*(11-12), 19-40.

Jakobson, R. (1968). Child language, aphasia and phonological universals. In *Child Language, Aphasia and Phonological Universals*. De Gruyter Mouton.

Javed, K., Reddy, V., & Lui, F. (2021). Neuroanatomy, choroid plexus. In *StatPearls [Internet]*. StatPearls Publishing.

Jessen, F., Erb, M., Klose, U., Lotze, M., Grodd, W., & Heun, R. (1999). Activation of human language processing brain regions after the presentation of random letter strings demonstrated with event-related functional magnetic resonance imaging. *Neuroscience Letters*, *270*(1), 13-16.

Johnson, M. H. (2000). Functional brain development in infants: elements of an interactive specialization framework. *Child Dev*, *71*(1), 75-81. https://doi.org/10.1111/1467-8624.00120

Jürgens, U. (1998). Speech evolved from vocalization, not mastication. *Behavioral and Brain Sciences*, *21*(4), 519-520.

Jürgens, U. (2003). From mouth to mouth and hand to hand: On language evolution. *Behavioral and Brain Sciences*, *26*(2), 229-230.

Jusczyk, P. W. & Aslin, R. N. (1995). Infants’ detection of the sound patterns of words in fluent speech. Cognitive Psychology, 29, 1–23.

Kaiser, A., Kuenzli, E., Zappatore, D., & Nitsch, C. (2007). On females' lateral and males' bilateral activation during language production: A fMRI study. *International Journal of Psychophysiology*, *63*(2), 192-198. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2006.03.008

Kaneko, N., Sawada, M., & Sawamoto, K. (2017). Mechanisms of neuronal migration in the adult brain. *J Neurochem*, *141*(6), 835-847. https://doi.org/10.1111/jnc.14002

Keenan, J. P., Thangaraj, V., Halpern, A. R., & Schlaug, G. (2001). Absolute pitch and planum temporale. *NeuroImage*, *14*(6), 1402-1408.

Kent, R. D. (1984). Psychobiology of speech development: coemergence of language and a movement system. *Am J Physiol*, *246*(6 Pt 2), R888-894. https://doi.org/10.1152/ajpregu.1984.246.6.R888

Kent, R. D. (1993). Infants and speech: Seeking patterns. *Journal of Phonetics*, *21*(1-2), 117-123.

Kent, R. D. (2015). Nonspeech Oral Movements and Oral Motor Disorders: A Narrative Review. *Am J Speech Lang Pathol*, *24*(4), 763-789. https://doi.org/10.1044/2015_ajslp-14-0179

Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: A review.

Keren-Portnoy, T., DePaolis, R. A., & Vihman, M. M. (2005). The articulatory filter and the creation of sound-meaning links: A developmental study with implications for evolution. Emergence of Linguistic Abilities (ELA) conference, Lyon, France,

Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J., & Williams, N. M. (2010). The role of vocal practice in constructing phonological working memory.

Keren-Portnoy, T., Daffern, H., Depaolis, R. A., Cox, C. M. M., Brown, K. I., Oxley, F. A. R., & Kanaan, M. (2021). “Did I just do that?”—Six-month-olds learn the contingency between their vocalizations and a visual reward in 5 minutes. *Infancy*, *26*(6), 1057-1075. https://doi.org/10.1111/infa.12433

Kern, M., Bert, S., Glanz, O., Schulze-Bonhage, A., & Ball, T. (2019). Human motor cortex relies on sparse and action-specific activation during laughing, smiling and speech production. *Communications Biology*, *2*(1), 118. https://doi.org/10.1038/s42003-019-0360-3

King, B. J., & Shanker, S. G. (2003). How can we know the dancer from the dance? *Anthropological Theory*, *3*(1), 5-26. https://doi.org/10.1177/1463499603003001749

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein, E.-B., & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12), 2512-2518. https://doi.org/10.1093/brain/123.12.2512

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (Saccopteryx bilineata). *Naturwissenschaften*, *93*(9), 451-454.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: an fMRI study. *NeuroImage*, *25*(4), 1068-1076.

Koopmans-Van Beinum, F. J., Clement, C. J., & Van Den Dikkenberg-Pot, I. (2001). Babbling and the lack of auditory speech perception: a matter of coordination? *Developmental Science*, *4*(1), 61-70. https://doi.org/10.1111/1467-7687.00149

Kovach, C. K., Daw, N. D., Rudrauf, D., Tranel, D., O'Doherty, J. P., & Adolphs, R. (2012). Anterior prefrontal cortex contributes to action selection through tracking of recent reward trends. *Journal of Neuroscience*, *32*(25), 8434-8442.

Kunzetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: tests in linear mixed effect models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Laing, C. and Bergelson, E. 2020. From babble to words: Infants' early productions match words and objects in their environment. *Cognitive Psychology* 122, article number: 101308. https://doi.org/10.1016/j.cogpsych.2020.101308

Leavens, D. A. (2003). Integration of visual and vocal communication: Evidence for Miocene origins. *Behavioral and Brain Sciences*, *26*(2), 232-233.

Leech, R., Braga, R., & Sharp, D. J. (2012). Echoes of the brain within the posterior cingulate cortex. *Journal of Neuroscience*, *32*(1), 215-222.

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain*, *137*(1), 12-32.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends in Neurosciences*, *16*(11), 444-447. https://doi.org/https://doi.org/10.1016/0166-2236(93)90072-T

Lenhoff, H. M., Wang, P. P., Greenberg, F., & Bellugi, U. (1997). Williams syndrome and the brain. *Scientific American*, *277*(6), 68-73.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. John Wiley and Sons, Inc.

Lieberman, P. (2017). Comment on "Monkey vocal tracts are speech-ready". *Science advances*, *3*(7), e1700442-e1700442. https://doi.org/10.1126/sciadv.1700442

Lindblom, B. (1998). A curiously ubiquitous articulatory movement. *Behavioral and Brain Sciences*, *21*(4), 521-522.

Lindell, A. (2013). Continuities in Emotion Lateralization in Human and Non-Human Primates [Review]. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00464

Lindell, A. K. (2006). In Your Right Mind: Right Hemisphere Contributions to Language Processing and Production. *Neuropsychology Review*, *16*(3), 131-148. https://doi.org/10.1007/s11065-006-9011-9

Lindell, A. K. (2016). Atypical hemispheric asymmetry in fetal alcohol spectrum disorders: a review of the effects of prenatal alcohol exposure on language lateralization. *Acta Neuropsychologica*, *14*(4), 367-380. https://doi.org/DOI: 10.5604/17307503.1227531

Lindell, A. K. (2020). Does Atypical Lateralization Influence Comorbid Psychopathology in Children with Autism Spectrum Disorders? *Advances in Neurodevelopmental Disorders*, *4*(1), 85-96.

Lindell, A. K., Tenenbaum, H. R., & Aznar, A. (2017). Left cheek bias for emotion perception, but not expression, is established in children aged 3–7 years. *Laterality: Asymmetries of Body, Brain and Cognition*, *22*(1), 17-30. https://doi.org/10.1080/1357650x.2015.1108328

Locke, J. L. (1983). *Phonological Acquisition and Change*. Academic Press.

Locke, J. L. (2000). Movement Patterns in Spoken Language. *Science*, *288*(5465), 449-451. https://doi.org/doi:10.1126/science.288.5465.449

Locke, J. L., & Pearson, D. M. (1990). Linguistic significance of babbling: Evidence from a tracheostomized infant. *Journal of Child Language*, *17*(1), 1-16.

Locke, J. L., & Pearson, D. M. (1992). Vocal learning and the emergence of phonological capacity: A neurobiological approach. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, research, implications* (pp. 91-129). York Press.

Losin, E. A. R., Russell, J. L., Freeman, H., Meguerditchian, A., & Hopkins, W. D. (2008). Left Hemisphere Specialization for Oro-Facial Movements of Learned Vocal Signals by Captive Chimpanzees. *PLoS ONE*, *3*(6), e2529. https://doi.org/10.1371/journal.pone.0002529

Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social*

Science. R package version 2.8.10. https://CRAN.R-project.org/package=sjPlot

Luef, E. M. (2018). Tracing the human brain's classical language areas in extant and extinct hominids. *The talking*, 29.

Lund, J. P. (1998). Is speech just chewing the fat? *Behavioral and Brain Sciences*, *21*(4), 522-522.

Lund, J. P., & Kolta, A. (2006). Brainstem circuits that control mastication: do they have anything to say during speech? *Journal of communication disorders*, *39*(5), 381-390.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*(1), 41-74. https://doi.org/10.1017/S0305000900007029

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behav Brain Sci*, *21*(4), 499-511; discussion 511-446. https://doi.org/10.1017/s0140525x98001265

MacNeilage, P. F. (2003). Mouth to hand and back again? Could language have made those journeys? *Behavioral and Brain Sciences*, *26*(2), 233-234.

MacNeilage, P. F. (2008). *The Origin of Speech*. Oxford University Press.

MacNeilage, P. F., & Davis, B. L. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech, Language, and Hearing Research*, *33*(1), 16-27.

MacNeilage, P. F., & Davis, B. L. (1993). Motor explanations of babbling and early speech patterns. In B. d. Boysson-Bardies, S. d. Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 341-352). Springer.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, *288*(5465), 527-531.

MacNeilage, P. F., & Davis, B. L. (2001). The role of rhythmic cyclicities in infant action development. *Developmental Science*, *4*(1), 79-83.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Dev*, *71*(1), 153-163. https://doi.org/10.1111/1467-8624.00129

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., Grebe, R., & Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, *110*(12), 4846-4851. https://doi.org/10.1073/pnas.1212220110

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*(2), 179-204. https://doi.org/10.1080/15475441.2013.829740

Marcovitch, S., & Lewkowicz, D. J. (2004). U-Shaped Functions: Artifact or Hallmark of Development? *Journal of Cognition and Development*, *5*(1), 113-118. https://doi.org/10.1207/s15327647jcd0501_10

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C., & Westerman, G. (2007). *Neuroconstructivism: How the brain constructs cognition* (Vol. Volume One). Oxford University Press.

Marie, D., Roth, M., Lacoste, R., Nazarian, B., Bertello, A., Anton, J.-L., Hopkins, W. D., Margiotoudi, K., Love, S. A., & Meguerditchian, A. (2018). Left Brain Asymmetry of the

Planum Temporale in a Nonhominid Primate: Redefining the Origin of Brain Specialization for Language. *Cerebral Cortex*, *28*(5), 1808-1815. https://doi.org/10.1093/cercor/bhx096

Markze, M., & Markze, R. (2000). Evolution of the human hand: Approaches to acquiring, analysing and interpreting the anatomical evidence. *The Journal of Anatomy, 197*(1), 121-140. doi:10.1046/j.1469-7580.2000.19710121.x

Márquez, S., Mowbray, K., Sawyer, G. J., Jacob, T., & Silvers, A. (2001). New fossil hominid calvaria from Indonesia-Sambungmacan 3. *The Anatomical Record*, *262*(4), 344-368. https://doi.org/10.1002/ar.1046

Marshall, A. J., Wrangham, R. W., & Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour*, *58*(4), 825-830.

Masataka, N. (2001). Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate–limiting constraint on the emergence of canonical babbling. *Developmental Science*, *4*(2), 158-164.

Masataka, N., & Fujita, K. (1989). Vocal learning of Japanese and rhesus monkeys. *Behaviour*, *109*(3-4), 191-199.

May, L., Byers-Heinlein, K., Gervain, J., & Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Front Psychol*, *2*, 222. https://doi.org/10.3389/fpsyg.2011.00222

McCune, L. (1998). Frame dominance: A developmental phenomenon? *Behavioral and Brain Sciences*, *21*(4), 522-523.

McCune, L., & Vihman, M. M. (1987). Vocal Motor Schemes. *Papers and Reports on Child Language Development*, *26*, 72-79.

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development: a productivity approach. *J Speech Lang Hear Res*, *44*(3), 670-684. https://doi.org/10.1044/1092-4388(2001/054)

McManus, I. C., Sik, G., Cole, D. R., Mellon, A. F., Wong, J., & Kloss, J. (1988). The development of handedness in children. *British Journal of Developmental Psychology*, *6*(3), 257-273.

Meier, R. P., McGarvin, L., Zakia, R. A., & Willerman, R. (1997). Silent mandibular oscillations in vocal babbling. *Phonetica*, *54*(3-4), 153-171. https://doi.org/10.1159/000262219

Meins, E. (1998). The effects of security of attachment and maternal attribution of meaning on children's linguistic acquisitional style. *Infant Behavior & Development*, *21*(2), 237-252. https://doi.org/10.1016/S0163-6383(98)90004-2

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*(19), 1692-1696.

Menn, L. (1971). Phonotactic rules in beginning speech: a study in the development of English discourse. *Lingua*, *26*(3), 225-251.

Menn, L. (1998). A multi-modal, emergent view of the development of syllables in early phonology. *Behavioral and Brain Sciences*, *21*(4), 523-524.

Mervis, C. B., & Becerra, A. M. (2007). Language and communicative development in Williams syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 3-15. https://doi.org/10.1002/mrdd.20140

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science*, *343*(6174), 1006-1010. https://doi.org/10.1126/science.1245994

Meyer, M., Elmer, S., & Jäncke, L. (2012). Musical expertise induces neuroplasticity of the planum temporale. *Annals of the New York Academy of Sciences*, *1252*(1), 116-123.

Michel, G. F. (2003). Ontogenetic constraints on the evolution of right-handedness. *Behavioral and Brain Sciences*, *26*(2), 234-235.

Millar, W. S. (1990). Span of Integration for Delayed-Reward Contingency Learning in 6- to 8-Month-Old Infants. *Annals of the New York Academy of Sciences*, *608*(1 The Developme), 239-266. https://doi.org/10.1111/j.1749-6632.1990.tb48899.x

Miller, J. L., & Kang, S. M. (2007). Preliminary ultrasound observation of lingual movement patterns during nutritive versus non-nutritive sucking in a premature infant. *Dysphagia*, *22*(2), 150-160.

Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: a developmental scenario. *Dev Cogn Neurosci*, *1*(3), 217-232. https://doi.org/10.1016/j.dcn.2011.03.005

Mitchell, P. R., & Kent, R. D. (1990). Phonetic variation in multisyllable babbling. *Journal of Child Language*, *17*(2), 247-265.

Monsalves, M. J., Bangdiwala, A. S., Thabane, A., & Bangdiwala, S. I. (2020). LEVEL (Logical Explanations & Visualizations of Estimates in Linear mixed models): recommendations for reporting multilevel data and analyses. *BMC Medical Research Methodology*, *20*(1), 1-9.

Moore, C. A., & Ruark, J. L. (1996). Does speech emerge from earlier appearing oral motor behaviors? *Journal of Speech, Language, and Hearing Research*, *39*(5), 1034-1047.

Morillon, B., Lehongre, K., Frackowiak, R. S., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc Natl Acad Sci U S A*, *107*(43), 18688-18693. https://doi.org/10.1073/pnas.1007189107

Morrill, R. J., Paukner, A., Ferrari, P. F., & Ghazanfar, A. A. (2012). Monkey lipsmacking develops like the human speech rhythm. *Developmental Science*, *15*(4), 557-568. https://doi.org/10.1111/j.1467-7687.2012.01149.x

Müller, R. A., & Basho, S. (2004). Are nonlinguistic functions in "Broca's area" prerequisites for language acquisition? FMRI findings from an ontogenetic viewpoint. *Brain Lang*, *89*(2), 329-336. https://doi.org/10.1016/s0093-934x(03)00346-8

Mundy, P., Card, J., & Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *36*(4), 325-338.

Munhall, K. G., & Jones, J. A. (1998). Articulatory evidence for syllabic structure. *Behavioral and Brain Sciences*, *21*(4), 524-525.

Nagy, E. (2012). From symmetry to asymmetry? The development of smile. *Cortex*, *48*(8), 1064-1067.

Nathani Iyer, S., & Oller, D. K. (2008). Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss. *The Volta Review*, *108*(2), 115.

National Institute of Mental Health (NIMH). (2022, March). *Autism spectrum disorder*. Www.nimh.nih.gov. https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd

Nenert, R., Allendorfer, J. B., Martin, A. M., Banks, C., Vannest, J., Holland, S. K., & Szaflarski, J. P. (2017). Age-related language lateralization assessed by fMRI: The effects of sex and handedness. *Brain Res*, *1674*, 20-35. https://doi.org/10.1016/j.brainres.2017.08.021

Neubauer, S., Gunz, P., Scott, N. A., Hublin, J. J., & Mitteroecker, P. (2020). Evolution of brain lateralization: A shared hominid pattern of endocranial asymmetry is much more variable in humans than in great apes. *Science advances*, *6*(7), eaax9935. https://doi.org/10.1126/sciadv.aax9935

Newman, A. J., Supalla, T., Fernandez, N., Newport, E. L., & Bavelier, D. (2015). Neural systems supporting linguistic structure, linguistic experience, and symbolic communication in sign language and gesture. *Proceedings of the National Academy of Sciences*, *112*(37), 11684-11689. https://doi.org/doi:10.1073/pnas.1510527112

NHS. (2017, October 20). *Tracheostomy*. NHS. https://www.nhs.uk/conditions/tracheostomy

NHS. (2019, October 21). *Overview - Down's syndrome*. NHS. https://www.nhs.uk/conditions/downs-syndrome/

Nicholls, M., Orr, C., & Lindell, A. (2005). Magical ideation and its relation to lateral preference. *Laterality: Asymmetries of Body, Brain, and Cognition*, *10*(6), 503-515.

Nip, I. S., Green, J. R., & Marx, D. B. (2009). Early speech motor development: Cognitive and linguistic considerations. *J Commun Disord*, *42*(4), 286-298. https://doi.org/10.1016/j.jcomdis.2009.03.008

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: from action to language. *Physiology (Bethesda)*, *20*, 60-69. https://doi.org/10.1152/physiol.00043.2004

Ocklenburg, S., Beste, C., Arning, L., Peterburs, J., & Güntürkün, O. (2014). The ontogenesis of language lateralization and its relation to handedness. *Neurosci Biobehav Rev*, *43*, 191-198. https://doi.org/10.1016/j.neubiorev.2014.04.008

Ohala, J. J. (1998). Content first, frame later. *Behavioral and Brain Sciences*, *21*(4), 525-526.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 1, pp. 93-112). Academic Press.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Dev*, *59*(2), 441-449. https://www.ncbi.nlm.nih.gov/pubmed/3359864

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science*, *4*(1), 49-60. https://doi.org/10.1111/1467-7687.00148

Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., & Call, J. (2019). Language origins viewed in spontaneous and interactive vocal rates of human and bonobo infants. *Frontiers in Psychology*, *10*, 729.

Oller, D. K., Wieman, L. A., Doyle, W., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language*, *3*, 1-11.

Oller, D. K. a. (2000). *The emergence of the speech capacity*. Erlbaum.

Olulade, O. A., Seydell-Greenwald, A., Chambers, C. E., Turkeltaub, P. E., Dromerick, A. W., Berl, M. M., Gaillard, W. D., & Newport, E. L. (2020). The neural basis of language development: Changes in lateralization over age. *Proceedings of the National Academy of Sciences*, *117*(38), 23477-23483. https://doi.org/doi:10.1073/pnas.1905590117

Ono, K., Nakamura, A., Yoshiyama, K., Kinkori, T., Bundo, M., Kato, T., & Ito, K. (2011). The effect of musical experience on hemispheric lateralization in musical feature processing. *Neuroscience Letters*, *496*(2), 141-145.

Onozuka, M., Fujita, M., Watanabe, K., Hirano, Y., Niwa, M., Nishiyama, K., & Saito, S. (2002). Mapping brain region activity during chewing: a functional magnetic resonance imaging study. *Journal of Dental Research*, *81*(11), 743-746.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, *106*(51), 22026-22031.

Oxley, F. A. R., Keren-Portnoy, T., Zweig, E., & Vihman, M. M. (2014). *Does longitudinal analysis of mouth aperture asymmetry in canonical babble indicate growing left hemisphere specialisation?* University of York].

Paquette, N., Lassonde, M., Vannasing, P., Tremblay, J., González-Frankenberger, B., Florea, O., Béland, R., Lepore, F., & Gallagher, A. (2015). Developmental patterns of expressive language hemispheric lateralization in children, adolescents and adults using functional near-infrared spectroscopy. *Neuropsychologia*, *68*, 117-125. https://doi.org/10.1016/j.neuropsychologia.2015.01.007

Paredes, M. F., James, D., Gil-Perotin, S., Kim, H., Cotter, J. A., Ng, C., Sandoval, K., Rowitch, D. H., Xu, D., McQuillen, P. S., Garcia-Verdugo, J.-M., Huang, E. J., & Alvarez-Buylla, A. (2016). Extensive migration of young neurons into the infant human frontal lobe. *Science*, *354*(6308), aaf7073. https://doi.org/doi:10.1126/science.aaf7073

Parra-López, P., Olmos-Soria, M., & Valero-García, A. V. (2022). Nonverbal Oro-Motor Exercises: Do They Really Work for Phonoarticulatory Difficulties? *International*

Journal of Environmental Research and Public Health, *19*(9), 5459. https://doi.org/10.3390/ijerph19095459

Patel, S., Oishi, K., Wright, A., Sutherland-Foggio, H., Saxena, S., Sheppard, S. M., & Hillis, A. E. (2018). Right hemisphere regions critical for expression of emotion through prosody. *Frontiers in neurology*, *9*, 224.

Paterson, S. J., Heim, S., Friedman, J. T., Choudhury, N., & Benasich, A. A. (2006). Development of structure and function in the infant brain: implications for cognition, language and social behaviour. *Neurosci Biobehav Rev*, *30*(8), 1087-1105. https://doi.org/10.1016/j.neubiorev.2006.05.001

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of autism and developmental disorders*, *44*(10), 2413-2428.

Payne, H., Gutierrez-Sigut, E., Woll, B., & MacSweeney, M. (2019). Cerebral lateralisation during signed and spoken language production in children born deaf. *Dev Cogn Neurosci*, *36*, 100619. https://doi.org/10.1016/j.dcn.2019.100619

Pedersen, A. V., & Vereijken, B. (2003). Laterality probabilities fluctuate during ontogenetic development. *Behavioral and Brain Sciences*, *26*(2), 236-237.

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., & Friederici, A. D. (2011). Neural language networks at birth. *Proc Natl Acad Sci U S A*, *108*(38), 16056-16061. https://doi.org/10.1073/pnas.1102991108

Perret, M. (1986). Social influences on oestrous cycle length and plasma progesterone concentrations in the female lesser mouse lemur (Microcebus murinus). *Reproduction*, *77*(1), 303-311.

Persson, A. E., Al-Khatib, D., & Flynn, T. (2020). Hearing aid use, auditory development, and auditory functional performance in swedish children with moderate hearing loss during the first 3 years. *American Journal of Audiology*, 1-14.

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M. R., Uylings, H. B. M., Rakic, P., & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences*, *108*(32), 13281-13286. https://doi.org/doi:10.1073/pnas.1105108108

Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B., & Ostry, D. J. (2004). Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on the hands. *Cognition*, *93*(1), 43-73. https://doi.org/10.1016/j.cognition.2003.10.007

Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science*, *251*(5000), 1493-1496. https://doi.org/10.1126/science.2006424

Petrides, M., & Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biology*, *7*(8), e1000170.

Poremba, A. (2006). Auditory processing and hemispheric specialization in non-human primates. *Cortex*, *42*(1), 87-89. https://doi.org/10.1016/s0010-9452(08)70325-3

Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L., & Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science*, *299*(5606), 568-572.

Previc, F. H. (1991). A general theory concerning the prenatal origins of cerebral lateralization in humans. *Psychol Rev*, *98*(3), 299-334. https://doi.org/10.1037/0033-295x.98.3.299

Priestly, T. M. S. (1977). One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language*, *4*(1), 45-65. https://doi.org/10.1017/s0305000900000477

Provins, K. A. (1997). Handedness and speech: a critical reappraisal of the role of genetic and environmental factors in the cerebral lateralization of function. *Psychol Rev*, *104*(3), 554-571. https://doi.org/10.1037/0033-295x.104.3.554

Purkayastha, S., & Sorond, F. (2013). Transcranial Doppler Ultrasound: Technique and Application. *Seminars in Neurology*, *32*(04), 411-420. https://doi.org/10.1055/s-0032-1331812

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal hearing. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *28*(3), 191-212. https://doi.org/https://doi.org/10.1016/0028-2243(88)90030-5

R Development Core Team. (2021). *R: A language and environment for statistical computing.* In R Foundation for Statistical Computing. http://www.R-project.org

Raja Beharelle, A., Dick, A. S., Josse, G., Solodkin, A., Huttenlocher, P. R., Levine, S. C., & Small, S. L. (2010). Left hemisphere regions are critical for language in the face of early left focal brain injury. *Brain*, *133*(6), 1707-1716.

Ramsay, D. S. (1980). Beginnings of bimanual handedness and speech in infants. *Infant Behavior and Development*, *3*, 67-77. https://doi.org/https://doi.org/10.1016/S0163-6383(80)80007-5

Ramsay, D. S. (1984). Onset of duplicated syllable babbling and unimanual handedness in infancy: Evidence for developmental change in hemispheric specialization? *Developmental Psychology*, *20*(1), 64-71. https://doi.org/10.1037/0012-1649.20.1.64

Rasmussen, T., & Milner, B. (1977). The Role of Early Left-Brain Injury in Determining Lateralization of Cerebral Speech Functions. *Annals of the New York Academy of Sciences*, *299*(1 Evolution and), 355-369. https://doi.org/10.1111/j.1749-6632.1977.tb41921.x

Reissland, N., Francis, B., Aydin, E., Mason, J., & Exley, K. (2014). Development of prenatal lateralization: evidence from fetal mouth movements. *Physiology & behavior*, *131*, 160-163.

Riès, S. K., Dronkers, N. F., & Knight, R. T. (2016). Choosing words: left hemisphere, right hemisphere, or both? Perspective on the lateralization of word retrieval. *Ann N Y Acad Sci*, *1369*(1), 111-131. https://doi.org/10.1111/nyas.12993

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin*, *95*(1), 52-77. https://doi.org/10.1037/0033-2909.95.1.52

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*(1), 169-192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nature Reviews Neuroscience*, *17*(12), 757-765. https://doi.org/10.1038/nrn.2016.135

Roach, N. T., & Richmond, B. G. (2015). Clavicle length, throwing performance and the reconstruction of the Homo erectus shoulder. *Journal of human evolution*, *80*, 107-113.

Rosenberg, K., & Trevathan, W. (2005). Bipedalism and human birth: The obstetrical dilemma revisited. *Evolutionary Anthropology: Issues, News, and Reviews*, *4*(5), 161-168. https://doi.org/10.1002/evan.1360040506

Rosselli, M., Ardila, A., Matute, E., & Vélez-Uribe, I. (2014). Language Development across the Life Span: A Neuropsychological/Neuroimaging Perspective. *Neuroscience Journal*, *2014*, 1-21. https://doi.org/10.1155/2014/585237

Rothbart, M. K., Taylor, S. B., & Tucker, D. M. (1989). Right-sided facial asymmetry in infant emotional expression. *Neuropsychologia*, *27*(5), 675-687. https://doi.org/10.1016/0028-3932(89)90112-7

Rovee-Collier, C. (1997). Dissociations in infant memory: rethinking the development of implicit and explicit memory. *Psychol Rev*, *104*(3), 467-498. https://doi.org/10.1037/0033-295x.104.3.467

Rudis, B. (2020). *hrbrthemes: Additional Themes, Theme Components and Utilities for 'ggplot2'. R package version 0.8.0.* https://CRAN.R-project.org/package=hrbrthemes

Russell, J. L., McIntyre, J. M., Hopkins, W. D., & Taglialatela, J. P. (2013). Vocal learning of a communicative signal in captive chimpanzees, Pan troglodytes. *Brain and Language*, *127*(3), 520-525.

Saarinen, T., Laaksonen, H., Parviainen, T., & Salmelin, R. (2005). Motor Cortex Dynamics in Visuomotor Production of Speech and Non-speech Mouth Movements. *Cerebral Cortex*, *16*(2), 212-222. https://doi.org/10.1093/cercor/bhi099

Salmelin, R., & Sams, M. (2002). Motor cortex involvement during verbal versus non-verbal lip and tongue movements. *Human brain mapping*, *16*(2), 81-91.

Santamaria, L., Noreika, V., Georgieva, S., Clackson, K., Wass, S., & Leong, V. (2020). Emotional valence modulates the topology of the parent-infant inter-brain network. *NeuroImage*, *207*, 116341. https://doi.org/https://doi.org/10.1016/j.neuroimage.2019.116341

Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., & Slocombe, K. E. (2013a). Chimpanzee food calls are directed at specific individuals. *Animal Behaviour*, *86*(5), 955-965.

Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013b). Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality. *PLoS ONE*, *8*(10), e76674. https://doi.org/10.1371/journal.pone.0076674

Scheumann, M., & Zimmermann, E. (2008). Sex-specific asymmetries in communication sound perception are not related to hand preference in an early primate. *BMC Biology*, *6*(1), 3. https://doi.org/10.1186/1741-7007-6-3

Schmitz, J., Kumsta, R., Moser, D., Güntürkün, O., & Ocklenburg, S. (2018). DNA methylation in candidate genes for handedness predicts handedness direction. *Laterality: Asymmetries of Body, Brain and Cognition*, *23*(4), 441-461. https://doi.org/10.1080/1357650x.2017.1377726

Schuetze, P., & Reid, H. M. (2005). Emotional lateralisation in the second year of life: evidence from oral asymmetries. *Laterality*, *10*(3), 207-217. https://doi.org/10.1080/13576500442000030

Schulter-Ellis, F. P. (1980). Evidence of handedness on documented skeletons. *J Forensic Sci*, *25*(3), 624-630.

Serrien, D. J., Ivry, R. B., & Swinnen, S. P. (2006). Dynamics of hemispheric specialization and integration in the context of motor control. *Nat Rev Neurosci*, *7*(2), 160-166. https://doi.org/10.1038/nrn1849

Sessle, B. J. (1998). Recent evidence of the involvement of lateral frontal cortex in primate cyclic ingestive movements. *Behavioral and Brain Sciences*, *21*(4), 529-530.

Shott, S. R. (2006). Down syndrome: Common otolaryngologic manifestations. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *142C*(3), 131-140. https://doi.org/10.1002/ajmg.c.30095

Shumway-Cook, A. (2007). Vestibular Rehabilitation - An Effective, Evidence-Based Treatment. *Vestibular Disorders Association*. https://www.vestibular.org

Siniscalchi, M., D'Ingeo, S., Fornelli, S., & Quaranta, A. (2018). Lateralized behavior and cardiac activity of dogs in response to human emotional vocalizations. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-017-18417-4

Siniscalchi, M., Laddago, S., & Quaranta, A. (2016). Auditory lateralization of conspecific and heterospecific vocalizations in cats. *Laterality: Asymmetries of Body, Brain and Cognition*, *21*(3), 215-227. https://doi.org/10.1080/1357650x.2015.1116541

Siniscalchi, M., Quaranta, A., & Rogers, L. J. (2008). Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS ONE*, *3*(10), e3349.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, *101*(3), 260-277. https://doi.org/https://doi.org/10.1016/j.bandl.2007.02.008

Slocombe, K. E., Kaller, T., Turman, L., Townsend, S. W., Papworth, S., Squibbs, P., & Zuberbühler, K. (2010b). Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behavioral Ecology and Sociobiology*, *64*(12), 1959-1966. https://doi.org/10.1007/s00265-010-1006-0

Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, *104*(43), 17228-17233. https://doi.org/10.1073/pnas.0706741104

Smith, L. B., & Thelen, E. S. (1993). *A Dynamic systems approach to development : applications.* MIT Press.

Snowdon, C. T. (2017). Learning from monkey talk. *Science*, *355*(6330), 1120-1122. https://doi.org/doi:10.1126/science.aam7443

Sommer, I. E., Aleman, A., Somers, M., Boks, M. P., & Kahn, R. S. (2008). Sex differences in handedness, asymmetry of the Planum Temporale and functional language lateralization. *Brain Research*, *1206*, 76-88. https://doi.org/https://doi.org/10.1016/j.brainres.2008.01.003

Sommer, I. E., & Kahn, R. S. (2003). The left hemisphere as the redundant hemisphere. *Behavioral and Brain Sciences*, *26*(2), 239-240.

Sommer, I. E. C., Aleman, A., Bouma, A., & Kahn, R. S. (2004). Do women really have more bilateral language representation than men? A meta-analysis of functional imaging studies. *Brain*, *127*(8), 1845-1852. https://doi.org/10.1093/brain/awh207

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. (2006). Converging language streams in the human temporal lobe. *Journal of Neuroscience*, *26*(28), 7328-7336.

Spivey, M. (2008). *The continuity of mind*. Oxford University Press.

Stark. (1980). Stages of Speech Development in the First Year of Life. In G. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 1, pp. 73-92). Academic Press.

Starkweather, J. (2010). *Linear mixed effects modelling using r*. https://it.unt.edu/sites/default/files/linearmixedmodels_jds_dec2010.pdf

Steeve, R. W. (2010). Babbling and chewing: Jaw kinematics from 8 to 22 months. *Journal of Phonetics*, *38*(3), 445-458. https://doi.org/10.1016/j.wocn.2010.05.001

Steeve, R. W., Moore, C. A., Green, J. R., Reilly, K. J., & McMurtrey, J. R. (2008). Babbling, chewing, and sucking: Oromandibular coordination at 9 months.

Stoel-Gammon, C. (1992). Prelinguistic Vocal Development: Measurement and Predictions. In C. A. Fergusion, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 439-456). York Press.

Stoel-Gammon, C. (2001). Down syndrome phonology: Developmental patterns and intervention strategies. *Down syndrome research and practice*, *7*(3), 93-100.

Streri, A., & de Hevia, M. D. (2014). Manual lateralization in infancy. *Frontiers in Psychology*, *5*, 1575. https://doi.org/10.3389/fpsyg.2014.01575

Studdert-Kennedy, M. (1990). Language Development from an Evolutionary Perspective. *Haskins Laboratories Status Report on Speech Research*, *SR-101*(102), 14-27. https://doi.org/10.1.1.472.6346

Su, P., Kuan, C. C., Kaga, K., Sano, M., & Mima, K. (2008). Myelination progression in language-correlated regions in brain of normal children determined by quantitative MRI assessment. *Int J Pediatr Otorhinolaryngol*, *72*(12), 1751-1763. https://doi.org/10.1016/j.ijporl.2008.05.017

Sugiura, L., Toyota, T., Matsuba-Kurita, H., Iwayama, Y., Mazuka, R., Yoshikawa, T., & Hagiwara, H. (2017). Age-Dependent Effects of Catechol-O-Methyltransferase (COMT) Gene Val158Met Polymorphism on Language Function in Developing Children. *Cerebral Cortex*, *27*(1), 104-116. https://doi.org/10.1093/cercor/bhw371

Sussman, H. M. (2015). Why the Left Hemisphere Is Dominant for Speech Production: Connecting the Dots. *Biolinguistics*, *9*, 116-131. https://doi.org/10.5964/bioling.9035

Szaflarski, J. P., Allendorfer, J. B., Byars, A. W., Vannest, J., Dietz, A., Hernando, K. A., & Holland, S. K. (2014). Age at stroke determines post-stroke language lateralization. *Restorative Neurology and Neuroscience*, *32*(6), 733-742. https://doi.org/10.3233/rnn-140402

Taglialatela, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2008). Communicative Signaling Activates ‘Broca's’ Homolog in Chimpanzees. *Current Biology*, *18*(5), 343-348. https://doi.org/https://doi.org/10.1016/j.cub.2008.01.049

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, *349*(6249), 734-738. https://doi.org/doi:10.1126/science.aab1058

Telano, L. N., & Baker, S. (2021). Physiology, cerebral spinal fluid. In *StatPearls [Internet]*. StatPearls Publishing.

Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behaviour*, *27*, 699-715.

Thelen, E. (1981). Kicking, rocking, and waving: Contextual analysis of rhythmical stereotypies in normal human infants. *Animal Behaviour*, *29*(1), 3-11.

Thelen, E. (1989). Self-organization in developmental processes: Can systems approaches work? In *Systems and development.* (pp. 77-117). Lawrence Erlbaum Associates, Inc.

Thelen, E. (2005). Dynamic Systems Theory and the Complexity of Change. *Psychoanalytic Dialogues*, *15*(2), 255-283. https://doi.org/10.1080/10481881509348831

Thelen, E., & Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press. https://doi.org/10.7551/mitpress/2524.001.0001

Tierney, A. L., & Nelson, C. A., 3rd. (2009). Brain Development and the Role of Experience in the Early Years. *Zero Three*, *30*(2), 9-13.

Tobias, P. V. (1987). The brain of Homo habilis: A new level of organization in cerebral evolution. *Journal of human evolution*, *16*(7), 741-761. https://doi.org/https://doi.org/10.1016/0047-2484(87)90022-4

Trevathan, W. (2015). Primate pelvic anatomy and implications for birth. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1663), 20140065.

Vallone, S., & Carnegie-Hargreaves, F. (2016). The infant with dysfunctional feeding patterns–The chiropractic assessment. *JOURNAL OF CLINICAL CHIROPRACTIC PEDIATRICS*, *15*(2), 1230-1235.

Vallortigara, G., Rogers, L. J., & Bisazza, A. (1999). Possible evolutionary origins of cognitive brain lateralization. *Brain Res Brain Res Rev*, *30*(2), 164-175. https://doi.org/10.1016/s0165-0173(99)00012-0

Van Der Haegen, L., & Brysbaert, M. (2018). The relationship between behavioral language laterality, face laterality and language performance in left-handers. *PLoS ONE*, *13*(12), e0208696. https://doi.org/10.1371/journal.pone.0208696

Van Der Knaap, M. S., & Valk, J. (1990). MR imaging of the various stages of normal myelination during the first year of life. *Neuroradiology*, *31*(6), 459-470. https://doi.org/10.1007/bf00340123

Van der Stelt, J. M., & Koopmans van-Beinum, F. J. (1986). Early stages in the development of speech movements. In *Precursors of early speech* (pp. 37-50). Springer.

Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., Wallois, F., Lepore, F., Béland, R., Lassonde, M., & Gallagher, A. (2016). Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. *Neuropsychologia*, *84*, 63-69. https://doi.org/https://doi.org/10.1016/j.neuropsychologia.2016.01.038

Vernes, S. C., Janik, V. M., Fitch, W. T., & Slater, P. J. (2021). Vocal learning in animals and humans. In (Vol. 376, pp. 20200234): The Royal Society.

Vihman, 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). York Press.

Vihman, M., & Croft, W. (2007). Phonological development: Toward a "radical" templatic phonology. *Linguistics*, *45*(4), 683-725. https://doi.org/10.1515/LING.2007.021

Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, *21*(1-2), 61-82.

Vihman, M. M. (2002). The role of mirror neurons in the ontogeny of speech. *Mirror neurons and the evolution of brain and language*, 305-314.

Vihman, M. M. (2014). *Phonological development: the first two years* (Second edition. ed.). Wiley-Blackwell.

Vihman, M. M. (2019). *Phonological Templates in Development*. Oxford University Press.

Vihman, M. M., & DePaolis, R. A. (2000). The role of mimesis in infant language development: Evidence for phylogeny? In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form* (pp. 130-145). Cambridge University Press.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2009). Babbling and words: A Dynamic Systems perspective on phonological development. In E. Bavin (Ed.), *Handbook of Child Language* (pp. 163-182). Cambridge University Press.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2014). The role of production in infant word learning. *Language Learning*, *64*(Suppl 2), 121-140. https://doi.org/10.1111/lang.12058

Vihman, M. M., Kay, E., de Boysson-Bardies, B., Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-yr-old children. *Developmental Psychology*, *30*(5), 651-662. https://doi.org/10.1037/0012-1649.30.5.651

Vihman, M. M., Macken, M. A., Miller, R. L., Simmons, H., & Miller, J. L. (1985). From Babbling to Speech: A Re-Assessment of the Continuity Issue. *Language*, *61*, 397.

Vr ieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, *17*(2), 228.

Wada, J. A. (1977). Pre-language and fundamental asymmetry of the infant brain. *Ann N Y Acad Sci*, *299*, 370-379. https://doi.org/10.1111/j.1749-6632.1977.tb41922.x

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic bulletin & review*, *11*, 192-196

Walker, A., & Leakey, R. E. (1993). *The Nariokotome Homo Erectus Skeleton*. Harvard University Press.

Wallentin, M. (2018). Sex differences in post-stroke aphasia rates are caused by age. A meta-analysis and database query. *PLoS ONE*, *13*(12), e0209571. https://doi.org/10.1371/journal.pone.0209571

Wallez, C., Schaeffer, J., Meguerditchian, A., Vauclair, J., Schapiro, S. J., & Hopkins, W. D. (2012). Contrast of hemispheric lateralization for oro-facial movements between learned attention-getting sounds and species-typical vocalizations in chimpanzees: Extension in a second colony. *Brain and Language*, *123*(1), 75-79. https://doi.org/https://doi.org/10.1016/j.bandl.2012.07.002

Wallez, C., & Vauclair, J. (2011). Right hemisphere dominance for emotion processing in baboons. *Brain and Cognition*, *75*(2), 164-169.

Wallez, C., & Vauclair, J. (2012). First evidence of population-level oro-facial asymmetries during the production of distress calls by macaque (Macaca mulatta) and baboon (Papio anubis) infants. *Behav Brain Res*, *234*(1), 69-75. https://doi.org/10.1016/j.bbr.2012.06.004

Wang, J., Yamasaki, B. L., Weiss, Y., & Booth, J. R. (2021). Both frontal and temporal cortex exhibit phonological and semantic specialization during spoken language processing in

7- to 8-year-old children. *Hum Brain Mapp*, *42*(11), 3534-3546. https://doi.org/10.1002/hbm.25450

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological science*, *25*(7), 1314-1324.

Waterson, N. (1971). Child phonology: a prosodic view. *Journal of Linguistics*, *7*(2), 179-211. https://doi.org/10.1017/s0022226700002917

Watson, M. M., & Lof, G. L. (2008). Epilogue: what we know about nonspeech oral motor exercises. *Semin Speech Lang*, *29*(4), 339-344. https://doi.org/10.1055/s-0028-1103398

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., West, V., & Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, *25*(4), 495-499.

Weiss, Y., Cweigenberg, H. G., & Booth, J. R. (2018). Neural specialization of phonological and semantic processing in young children. *Hum Brain Mapp*, *39*(11), 4334-4348. https://doi.org/10.1002/hbm.24274

Whitworth, N., & Bray, M. (2015). Emergence of the vowel space in very young children with Down syndrome: An exploratory case study. ICPhS2015: Proceedings of the 18th International Congress of Phonetic Sciences, 10-14 August 2015, SECC, Glasgow, Scotland, UK,

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag.

Wickham, H., Averick, M., Bryan, J., Chang, W., Lucy D'Agostino McGowan, François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Spinu, D. P. S. V., Takahahi, K., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

William Syndrome Foundation (n.d.). Home page. https://william-syndrome.org.uk

Williamson, J. M., & Lyons, D. A. (2018). Myelin Dynamics Throughout Life: An Ever-Changing Landscape? [Mini Review]. *Frontiers in Cellular Neuroscience*, *12*. https://doi.org/10.3389/fncel.2018.00424

Wilson, E. M., Green, J. R., Yunusova, Y., & Moore, C. A. (2008). Task specificity in early oral motor development. Seminars in speech and language, Witelson, S. F., & Pallie, W. (1973). Left hemisphere specialization for language in the newborn. Neuroanatomical evidence of asymmetry. *Brain*, *96*(3), 641-646. https://doi.org/10.1093/brain/96.3.641

Wolf, M. E., & Goodale, M. A. (1987). Oral asymmetries during verbal and non-verbal movements of the mouth. *Neuropsychologia*, *25*(2), 375-396. https://doi.org/https://doi.org/10.1016/0028-3932(87)90026-1

Woll, B., & Sieratzki, J. S. (2003). Why homolaterality of language and hand dominance may not be the expression of a specific evolutionary link. *Behavioral and Brain Sciences*, *26*(2), 241-241.

Wolpert, L. (2003). Causal beliefs lead to toolmaking, which require handedness for motor control. *Behavioral and Brain Sciences*, *26*(2), 242-242.

Workman, L., Chilvers, L., Yeomans, H., & Taylor, S. (2006). Development of cerebral lateralisation for recognition of emotions in chimeric faces in children aged 5 to 11. *Laterality*, *11*(6), 493-507. https://doi.org/10.1080/13576500600724963

Wu, X., Holloway, R. L., Schepartz, L. A., & Xing, S. (2011a). A new brain endocast of Homo erectus from Hulu Cave, Nanjing, China. *American Journal of Physical Anthropology*, *145*(3), 452-460. https://doi.org/10.1002/ajpa.21527

Wu, X., & Pan, L. (2011b). Identification of Zhoukoudian Homo erectus brain asymmetry using 3D laser scanning. *Chinese Science Bulletin*, *56*(21), 2215-2220. https://doi.org/10.1007/s11434-011-4512-1

Wu, X., Schepartz, L. A., Falk, D., & Liu, W. (2006). Endocranial cast of Hexian Homo erectus from South China. *American Journal of Physical Anthropology*, *130*(4), 445-454. https://doi.org/10.1002/ajpa.20378

Wyler, F., Graves, R., & Landis, T. (1987). Cognitive task influence on relative hemispheric motor control: mouth asymmetry and lateral eye movements. *J Clin Exp Neuropsychol*, *9*(2), 105-116. https://doi.org/10.1080/01688638708405351

Wylie, D. R., & Goodale, M. A. (1988). Left-sided oral asymmetries in spontaneous but not posed smiles. *Neuropsychologia*, *26*(6), 823-832. https://doi.org/10.1016/0028-3932(88)90052-8

Yang, J., Asano, M., Kanazawa, S., Yamaguchi, M. K., & Imai, M. (2019). Sound symbolism processing is lateralized to the right temporal region in the prelinguistic infant brain. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-49917-0

Zhang, D., Zhou, Y., & Yuan, J. (2018). Speech Prosodies of Different Emotional Categories Activate Different Brain Regions in Adult Cortex: an fNIRS Study. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-017-18683-2

Zimmerman, E., & Foran, M. (2017). Patterned auditory stimulation and suck dynamics in full-term infants. *Acta Paediatrica*, *106*(5), 727-732. https://doi.org/10.1111/apa.13751

Zuberbühler, K. (2020). Syntax and compositionality in animal communication. *Philosophical Transactions of the Royal Society B*, *375*(1789), 20190062.