Individual differences in first language acquisition and their theoretical implications

Evan Kidd[123]

Amy Bidgood[4]

Seamus Donnelly[2]

Samantha Durrant[5]

Michelle S. Peter[5]

Caroline F. Rowland[156]

[1] Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2] Research School of Psychology, The Australian National University, Canberra, Australia

[3] ARC Centre of Excellence for the Dynamics of Language, Canberra, Australia

[4] School of Health and Society, University of Salford, Salford, UK

[5] ESRC LuCiD Centre, Department of Psychological Sciences, University of Liverpool, Liverpool, UK

[6] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Preface

In academia, we often talk in terms of academic families. Elena Lieven's academic family is huge, spanning generations; the oldest of the current authors first met her in 1992, the youngest in 2017, and all have benefitted from her mentorship. Elena taught her family many things, but three in particular stand out. First, she taught us that the scientific method matters; testing theories against what children actually say and understand is important. It is not legitimate to simply assume children have innate grammatical representations if there is no data to support this conclusion. Nor can we simply assume all children learn language in broadly similar ways – across individuals and cultures – if there is evidence for different developmental pathways. Second, pay attention to how language is used: children do not learn language in a vacuum but in a complex multi-modal environment, which contains a lot more information to language structure than you might think. And third, collaborate; because teams produce much stronger, much more exciting work than any one scientist ever could. Elena has never been a lone wolf researcher: one of her greatest strengths comes from her ability to learn from others, to discuss and hone her ideas in the company of others, to share her ideas generously, and to acknowledge and credit the influence of her collaborators on these ideas. We have all benefitted from this generosity of spirit and will always be grateful to her for it.

Abstract

Much of Lieven's pioneering work has helped move the study of individual differences to the centre of child language research. The goal of the present chapter is to illustrate how the study of individual differences provides crucial insights into the language acquisition process. In part one, we summarise some of the evidence showing how pervasive individual differences are across the whole of the language system; from gestures to morphosyntax. In part two, we describe three causal factors implicated in explaining individual differences, which, we argue, must be built into any theory of language acquisition (intrinsic differences in the neurocognitive learning mechanisms, the child's communicative environment, and developmental cascades in which each new linguistic skill that the child has to acquire depends critically on the prior acquisition of foundational abilities). In part three, we present an example study on the role of the speed of linguistic processing on vocabulary development, which illustrates our approach to individual differences. The results suggest a key role for the input in vocabulary acquisition, not only directly, by providing children with more opportunities for learning a greater diversity of words, but also indirectly, by increasing processing capacity, and thus speeding up the learning of new words. The results also show evidence of a changing relationship between lexical processing speed and vocabulary over developmental time, perhaps as a result of the changing nature of the structure of the lexicon. The study thus highlights the benefits of an individual differences approach in building, testing, and constraining theories of language acquisition.

Introduction

For many decades, language acquisition research was dominated by traditional formal linguistic approaches (e.g. Borer & Wexler, 1987; Hyams, 1986; Radford, 1990), which characterised individual differences in the trajectory of language acquisition as peripheral and even perhaps unimportant to explanations of language acquisition. The argument was that individual differences in the rate at which children learn words, or in the way language is processed by external cognitive systems, do not fundamentally affect the acquisition of the core representational properties of the linguistic system, which were considered to be the proper focus of research. According to many formal theories, since the range of variation in core representational properties is restricted by innate knowledge structures (so-called Universal Grammar), individual differences can exert very little effect on development, manifesting only in those neurocognitive disorders that result from genetically-caused abnormalities in brain development (e.g. in the left hemisphere and basal ganglia; van der Lely & Pinker, 2014).

Nowadays, thanks to the work of Elena Lieven and her contemporaries, colleagues and students, we have a better awareness of how linguistic systems (e.g. vocabulary and grammar) interact, a better understanding of how language processing can affect acquisition itself, and thus a better understanding of how and why individual differences manifest across linguistic domains. This has led to an increased interest in the role of individual differences and, crucially, their role in building, testing, and constraining our theories of language acquisition (see Kidd, Donnelly, & Christiansen, 2018, for a review).

This step-change in our attitude to individual differences can be traced back to over three decades of work in the emergentist tradition by usage-based (see e.g. Nelson, 1973; for a summary see Lieven, 2016) and constraint-based (e.g. Bates & MacWhinney, 1989) theorists. The central claim of these approaches is that children must induce knowledge about key properties of the language via analyses of, and generalization from, the input, with little language-specific prior knowledge - a process that places a large emphasis on the role of the input in acquisition and the learning mechanisms that process that input. Thus, they predict that individual differences in acquisition derive from two distinct sources: (i) variation in the intrinsic capacity of the neurocognitive learning mechanisms supporting language and (ii) variation in the richness of the communicative environment. The theories also assume a tight integration of form and meaning (i.e. syntax and semantics), which predicts meaningful interactions between levels of language that are also, themselves, subject to individual differences. The work testing these predictions has given us a large body of evidence on the role of variation in language development, particularly variation in the input, which has arguably influenced all modern theorising, from radical exemplar-based theories (Ambridge, 2018) to formal generativist models (e.g. Yang, Crain, Berwick, Chomsky, & Bolhuis, 2017). In other words, the legacy of the emergentist tradition has been to require all researchers to acknowledge individual differences, and to consider how they manifest in acquisition, what causes them, and what they mean for the acquisition process.

The goal of the present chapter is to illustrate how the study of individual differences provides crucial insights into the language acquisition process. In part one, we summarise some of the large body of evidence showing how pervasive individual differences are across the language system. In part two, we describe three causal factors implicated in explaining individual differences, factors which we argue must be built into any theory of language acquisition. In both parts we focus on research on monolingual (mainly English) acquisition, since this is where substantial work has been done, though note that there is a developing body of work on individual differences in multilingual populations (Hoff et al., 2012; Hoff, Quinn, & Giguere, 2018) and across languages (Frank, Braginsky, Marchman, & Yurovsky, n.d.). In part three, we present an example study that illustrates our approach to individual differences, and highlights the benefits of such an approach in building, testing, and constraining theories of language acquisition.

1. Individual differences in language acquisition

It is uncontroversial that there is significant variation in the trajectory of vocabulary acquisition not only in the number of words produced and comprehended by children at different ages, but also in the trajectory of growth, with some children growing their vocabulary much faster than others (see e.g. Fenson et al., 2007; Frank, Braginsky, Yurovsky, & Marchman, 2017). There is also increasing consensus that important individual differences exist across the entire linguistic system. These differences emerge early, manifesting first in children's earliest non-verbal communicative abilities. For example, McGillion et al. (2016) demonstrated large differences in the age of onset of both babble and pointing in a sample of 59 British English learning infants. Although all the infants had begun to babble by 15 months and to point using their index finger by 18 months, the age of onset of babble, defined as the stable production of two supra-glottal consonants (excluding glottal stops and glides) ranged from 9 to 15 months, and the range of onset of index finger pointing, defined according to criteria established in Matthews, Behne, Lieven, & Tomasello (2012), ranged from 9 to 18 months. This suggests continuity between early non-verbal and later language learning abilities, as well as suggesting that individual differences start to emerge early in life.

There are also substantial individual differences in the rate at which children acquire morphosyntax. Once again, individual differences manifest from the earliest stages; for example, some children are already using multi-word phrases and function words by 24 months, while others are still only producing single word utterances at that age (Bates, Bretheron, & Snyder, 1988). Individual variation is not restricted to the early stages though; data from Bishop's (2003) norming sample for the Test for the Reception of Grammar demonstrates large individual differences throughout childhood in the rate at which children acquire more complex grammatical constructions such as centre embedded clauses, relative clauses, passives and anaphor.

In addition, there are even individual differences in the strategies that children use to learn to produce multi-word utterances. Building on Bloom et al's. (1975) monograph demonstrating the existence of stylistic variation in children's early word-combinations, Pine and Lieven (1993) have suggested that there may be at least two separate routes into multiword speech. Some English-learning children seem to produce their earliest multiword utterances by combining two or more words from their single-word vocabularies together. The acquisition sequence here seems to be best characterised as a process of building up multi-word patterns from their constituent parts. However, other children seem to be breaking down, and reanalysing, multiword utterances originally learned as holistic phrases, a process that involves the child gaining productive control over ‘slots’ in the previously unanalysed phrases (e.g. *iwantdodat becomes I want X [do that/more/cookies/juice]*). The acquisition sequence here could, thus, be more accurately characterised as a process of developing patterns by varying words within initially unanalysed units. Note though that, although Bloom et al. characterised these two styles in terms of differences between children, Pine and Lieven have argued that all children use both routes, albeit to varying degrees.

Crucially, these individual differences are large and stable across development (Bates, Dale, & Thal, 1995; Bornstein & Putnick, 2012). Children who start late, and whose language development proceeds slowly in the first two years of life, are significantly more likely to reach language milestones late, and this has both theoretical and societal implications. In terms of theory development, establishing the causes of individual differences in language growth is central to our understanding of how language is acquired, and what factors (both intrinsic and extrinsic to the child) are implicated in this process. In societal terms, it is important because language skills at school entry affect a wide range of later developments; in one study of over 11,000 British children, those with poor vocabulary skills at age five were four times more likely to have reading difficulties in adulthood, three times as likely to have mental health problems and twice as likely to be unemployed when they reached adulthood (Law, Schoon, & Parsons, 2009). Determining the causes of slow language growth is crucial if we are to intervene effectively to improve the language skills of these children.

2. What causes individual differences in language acquisition?

Any complete theory of language acquisition must be able to explain what individual differences exist, and why such individual differences emerge in the developmental process. In this section we summarise work showing that (at least) three causal factors will ultimately need to be incorporated into our theories of language acquisition: intrinsic differences in neurocognitive learning mechanisms, the communicative environment including the linguistic input, and the role of developmental cascades.

2.1. Intrinsic differences in the neurocognitive learning mechanisms

Both our genetic inheritance and our early experiences (in the womb and in early infancy) influence early brain development, and thus fundamentally shape the brain's capacity for learning. Thus, different infants are likely to approach the language acquisition task with different intrinsic capacities for learning. Evidence for this is very clear from clinical populations, in which healthy brain development is disrupted, either because of genetic mutations or environmental triggers in pregnancy or early infancy. For example, we now have converging evidence for the role of the gene CNTNAP2 in the development of the circuits in the brain implicated in language acquisition (Abrahams et al., 2007). Variants of CNTNAP2 have been associated with “age at first word” in children with autism (Alarcón et al., 2008), reduced performance on standardised language tests in children with developmental language disorder (DLD, previously called specific language impairment, Vernes et al., 2008), and with susceptibility to autism and DLD within the general population (Whitehouse, Bishop, Ang, Pennell, & Fisher, 2011). Similarly, we also have good evidence for the role of some well-known environmental influences on early brain development. Maternal drug use, alcohol consumption, nutrition, and stress levels (as measured by cortisol) in pregnancy are all known to affect foetal brain development, causing adverse neurodevelopmental outcomes in later childhood (Bloomfield, 2011), and nutritional supplements can mitigate the risk of some of these developmental problems. For example, Roth et al. (2011) reported that maternal use of folic acid supplements in pregnancy, from 4 weeks before to 8 weeks after conception, was associated with a reduced risk of severe language delay at age 3 years in a sample of 44,420 children.

However, there are two caveats we must apply to the statements above. First, we do not yet know enough about the language circuits in the brain to understand *how* early genetic and environmental factors affect acquisition. It could be that they disrupt the child’s ability to access innate linguistic knowledge in some way (e.g. Rice, Wexler, & Cleave, 1995; van der Lely & Pinker, 2014). Alternatively, it could be that they disrupt the efficient functioning of more general cognitive learning mechanisms in ways that make language acquisition particularly difficult (e.g. causing deficits in verbal working memory or statistical learning abilities; Evans, Saffran, & Robe-Torres, 2009; Falcaro et al., 2008). These are all empirical issues that have yet to be answered by studies on brain development.

Second, we have only discussed studies of populations in which the disruption to brain functioning is so fundamental that it manifests as a developmental disorder. We do not yet know how these factors influence development within the normal range. Behavioural genetics studies suggest that there is a genetically inherited component to language ability in typically developing populations (Kovas et al., 2005) but these studies are limited to simply telling us *whether* there is a genetic component. They cannot provide, or test, explanatory models of *how* genes influence development, nor can they tell us much about the size of the genetic influence on behaviour, because it is difficult to generalise from the homogeneous Western, Educated, Industrialised, Rich and Democratic (WEIRD) samples in behavioural genetics studies to the general population (the more homogeneous the sample's environment, the more the heritability estimate will overestimate the amount of variance attributable to genetics; Johnson, Turkheimer, Gottesman, & Bouchard, 2010). It might be that the brain is incredibly resilient to small changes in the genetic code, such that these play very little role in later individual variation in language acquisition within the normal range. We thus need more work on the way in which genes, gene expression and the early environment work together to influence brain development at the cellular and cortical level in typically developing populations to determine how resilient the language acquisition process is to small changes in the development of the architecture of the brain.

2.2. The communicative environment (especially linguistic input)

Some readers may argue that the evidence presented in section 2.1 above is interesting but orthogonal to the key nature-nurture debate in language acquisition, which centres, much more narrowly, on the role of the child’s communicative environment in acquisition. But even if we describe environment more narrowly in terms of the communicative environment, there is a wealth of evidence for its effect on the trajectory of the language acquisition process.

Perhaps the most well researched environmental influence is the richness of the linguistic input that children receive in infancy and early childhood. Numerous studies have established a direct strong or medium correlation between measures of the quantity of linguistic input and the size of children’s vocabularies throughout development (Bornstein, Haynes, & Painter, 1998; Cartmill et al., 2013; Hart & Risley, 1995; Hoff & Naigles, 2002; Hurtado, Marchman, & Fernald, 2008; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2012; note that some of these studies also link variation in input quality and quantity with variation in socio-economic status but the basic underlying idea remains the same - that quality/quantity of input determines the size of a child's vocabulary and, by implication, the speed of vocabulary growth). For example, Hart and Risley’s (1992, 1995) data from American English learning children shows that some children (particularly those of high socio-economic status) were exposed to up to 153,000 more words per week than others, which had a significant effect on these children’s vocabulary growth. Pearson, Fernandez, Lewedeg, and Oller (1997) even showed that the relative vocabulary size of bilingual English-Spanish speaking one- to two-year-olds was predicted by the relative amount of input they received in each language, which suggests a direct association between the likelihood of hearing words in a language and the ability to learn them.

The quality of the linguistic input is also important. Lexical diversity, operationalized as the number of different word types produced by caregivers during a set timeframe, has a large to moderate, consistent, effect across a number of studies (Bornstein et al., 1998;

Demir-Vegter, Aarts, & Kurvers, 2014; Hoff & Naigles, 2002; Hsu, Hadley, & Rispoli, 2017; Pan, Rowe, Singer, & Snow, 2005). For example, Hoff and Naigles (2002) reported that the number of different word types produced by the mothers of 63 English-learning 2-year-olds was a strong predictor of the number of different words their children produced ten weeks later, and Weizman and Snow (2001) reported that older children benefit in particular from input that models a high proportion of rare words (see also Beals, 1997; Rowe, 2012). The complexity of input utterances, as measured by indices such as Mean Length of Utterance (MLU; Bornstein et al., 1998; Hoff & Naigles, 2002) and constituent and clausal complexity measures (Huttenlocher et al., 2010), also predict language growth. For example, Bornstein et al. (1998) used structural equation modelling to show that both maternal lexical diversity and maternal MLU were significant predictors of child vocabulary at 18 months. Similarly, the use of communicative devices that increase the chances that the child will interpret words correctly have been found to affect vocabulary growth. These include contingency (McGillion et al., 2013), referential (un)certainty (Cartmill et al., 2013), the number of utterances spoken during periods of joint attention (Hoff & Naigles, 2002), and how effectively parents model language during routines and rituals (Hirsh-Pasek et al., 2015). Thus, there is good evidence for a role for both the quantity of the input and its quality in determining the rate at which children develop vocabulary.

This effect of the communicative environment is not restricted to vocabulary acquisition. For example, in speech perception, there is evidence that mothers' vowel space size during speech directed at 6 to 12 month old infants correlated with their infants' performance in a speech perception task (Liu, Kuhl, & Tsao, 2003), and that differences in how distinctly caregivers distinguished between /s/ and /ʃ/ (as in *sip* /sɪp/ versus *ship* /ʃɪp/) predicted their infants' ability to discriminate the same sound pair beyond overall speech rate or pitch (Cristià, 2011). There is also evidence for effects of the input on morphosyntax acquisition (for a summary see Ambridge, Kidd, Rowland, & Theakston, 2015). This is a more controversial claim, given the traditional claim of formal generativist theory that the role of the input in the acquisition of the core representational properties of the linguistic system is minimal. However, these effects do exist and, as we are discovering, can be substantial. For example, Street and Dąbrowska (2014) have reported that adults with lower academic attainment were significantly slower, and made more errors when asked to identify the agent and patient roles in passive sentences, than those with higher academic attainment (see also Dąbrowska & Street, 2006; Street & Dabrowska, 2010). Although such explicit tasks call on other metalinguistic skills than just syntactic knowledge, which could be differentially available to adults with high and low levels of attainment, this was a relatively simple task, especially since participants were given a brief nontechnical explanation of the terms "do-er" and "acted-on" before the study started. Thus, the authors argued, since passives are more frequently encountered in written texts, the passive constructions of participants with more educational experience are better entrenched, and hence accessed more reliably, which results in faster and more accurate performance.

Once again, however, we must end the section on a caveat. As the late Judith Rich Harris reminded us (Harris, 1998), correlation is not cause in developmental psychology; in correlational studies such as many of those cited in this section above, it is always crucial to control for genetic correlations between parents and children (talkative parents have talkative children), and for the effects of the child on the adult (linguistically advanced children may elicit more, and more sophisticated, language from their caregivers). Modern work controlling for these factors demonstrates that there is still a role for such input factors after controlling for these confounds; for example, Romeo et al. (2018) have demonstrated that children who had experienced more conversational turns with adults exhibited greater left inferior frontal (Broca's area) activation, and that this significantly explained the relation between children's language exposure and verbal skill, even after controlling for both child and adult talkativeness. Regardless, we still need more work to map out how the child's developing neurocognitive mechanisms use the information in the environment to build linguistic knowledge; to understand how development is shaped by what Karmiloff-Smith calls the "the multiple two-way ... chains" of interaction between "the genetic, the brain in its spatial and temporal dynamics, the cognitive, the environmental and the behavioural" (Karmiloff-Smith, 1998: 397).

2.3. The role of developmental cascades

The phrase 'developmental cascades' refers to the "cumulative consequences for development of the many interactions and transactions occurring in developing systems that result in spreading effects across levels, among domains at the same level, and across different systems or generations" (Masten & Cicchetti, 2010: 491). Developmental cascades occur in language development because each new linguistic skill that the child has to acquire depends critically on the prior acquisition of foundational linguistic skills, on the child's existing language knowledge at the point of new learning, and on prior developments in related cognitive domains. In other words, knowledge states at time t determine how and what can subsequently be learned at time $t + 1$. Cascades affect every aspect of language acquisition and are likely to be responsible for the fact that small differences in foundational skills in infancy (e.g. in speech processing efficiency) can lead to considerable differences in learning rate, and thus substantial differences in vocabulary size by middle childhood.

Processing ability in early infancy seems to be an important predictor of later vocabulary development, both in infancy and in later childhood. For example, Brito, Fifer, Myers, Elliott, and Noble (2016) have reported that resting-state brain activity (low gamma

EEG power in the parietal cortex) in newborns correlated with language comprehension at 15 months (though not with expressive communication ability). Chonchaiya et al. (2013), measuring auditory brainstem responses to masked clicks, suggested that infants who show greater improvement in processing efficiency between 6 weeks and 9 months, have bigger vocabularies at 9 months of age (though note that processing efficiency at 6 months of age was not predictive of 9 month vocabulary, suggesting that the relationship develops during the first year of life). When we focus on skills more directly related to language, we find that both the ability to make phonetic distinctions and the ability to segment words from running speech are subject to significant inter-individual variation, and that children who master these skills early have better vocabulary than children who master them at an older age (Cristià, Seidl, Junge, Soderstrom, & Hagoort, 2014; Junge, Cutler, & Hagoort, 2012; Kidd, Junge, Spokes, Morrison, & Cutler, 2018; Tsao, Liu, & Kuhl, 2004). Infants' cortical auditory ERP responses, for rapidly presented auditory stimuli, pitch discriminations, and native vs. non-native speech contrasts are also related to later language skills (Choudhury, Leppanen, Leevers, & Benasich, 2007; Kuhl et al., 2008; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005).

There is also evidence for a role for early non-verbal communication skills in later language development. A number of studies have demonstrated that differences in early gesture use are robustly associated with later vocabulary growth, with early and frequent gesture users going onto develop larger vocabularies (see e.g. the meta-analysis by Colonnesi, Stams, Koster, & Noom, 2010). For example, as well as demonstrating individual differences in the age of pointing onset, McGillion et al. (2016) reported that age of onset of index finger pointing predicted receptive vocabulary size at 18 months (though note that it did not predict expressive vocabulary size).

More controversially perhaps, especially for theorists who characterise vocabulary and morphosyntactic development as driven by separate linguistic systems, morphosyntactic development seems to be reliant on vocabulary acquisition throughout early development (see Bates & Goodman, 1997). In fact, vocabulary and morphosyntactic development seem to be more tightly coupled than vocabulary production and vocabulary comprehension are. This is a controversial finding because, if vocabulary and morphosyntactic growth were governed by separate learning mechanisms, we would expect to see a tight coupling early on (as children need to know some words to get started on morphosyntactic development), but then we would expect the lines to diverge. Instead, the child's growth in vocabulary predicts their growth in grammar *throughout development*. This relationship also holds in every language for which we currently have data (Frank et al., n.d.).

Finally, there is evidence implicating a range of other cognitive abilities in language acquisition. Two clearly important precursor skills are phonological short-term memory and speed of online linguistic processing (for a discussion of verbal working memory, see Kidd, 2013). Variation in phonological short-term memory is measured by the non-word repetition (NWR) test in which children are asked to repeat back sequences of non-words. Although NWR is not a pure test of phonological short-term memory (performance on the task is strongly influenced by the size of the lexicon; see Jones, Gobet, & Pine, 2007), it is likely that some aspect of the variance can be attributed to phonological memory capacity (see e.g. Gathercole, 2006). NWR performance not only correlates with children's vocabulary and syntactic development (Caplan & Waters, 1999; Gathercole, Willis, Emslie, & Baddeley, 1992; Just & Carpenter, 1992), but can be used as a clinical marker for developmental language disorder (DLD) because it seems to discriminate well between typically developing children and children with DLD (Conti-Ramsden, Botting, & Faragher, 2001; Conti-Ramsden & Hesketh, 2003; Ellis Weismer et al., 2014, though see Stokes, Wong, Fletcher & Leonard, 2006). Similarly, the speed with which children can process linguistic information online in infancy (e.g. at 18 months) is a predictor of their later language development

(Fernald, Perfors, & Marchman, 2006), with effects on language acquisition being reported up to six years later (Marchman & Fernald, 2008).

Other cognitive processes are more domain general. Two that have been implicated in language acquisition are executive functions (EFs) and statistical learning (SL). EFs is an umbrella term for the management (regulation, control) of cognitive processes, including working memory, reasoning, attention switching, and problem solving, as well as planning and execution. EFs tend to be associated with explicit and sometimes effortful cognitive processes, and consequently have been linked to performance in linguistic tasks involving planning, regulation and/or choosing between competing presentations or responses, such as recovery from garden-path sentences (Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2014; Vuong & Martin, 2014; Woodard, Pozzan, & Trueswell, 2016), and competition in lexical processing (Festman, Rodriguez-Fornells, & Münte, 2010; Khanna & Boland, 2010; Nozari, Trueswell, & Thompson-Schill, 2016). SL refers to the general and pervasive tendency to identify and learn co-occurring elements in the environment. Since Saffran, Newport and Aslin (1996), we have known that children are capable of using statistical information to process language, but more recent evidence has suggested that humans vary in their SL capacity (e.g., Kaufman, DeYoung, Gray, Jimenez, Brown, & Mackintosh, 2010), which in children has been linked to natural language acquisition and use across a wide developmental range (e.g., infants: Hoareau, Yeung, & Nazzi, 2019; Kidd & Arciuli, 2016).

In sum, developmental cascades occur in language development because each new linguistic skill that the child has to acquire depends on the prior acquisition of foundational linguistic skills, on the child's existing language knowledge at the point of new learning, and on developments in related cognitive domains. However, again, we finish the section with two unresolved issues. First, establishing the causal mechanisms that underpin the relationship between foundational skills and later vocabulary acquisition is easier for some skills than others. For example, the reason why early speech perception abilities affect later vocabulary seems clear; the easier it is for a child to process auditory input, the easier it is for them to learn language from such input. However, the reason why early gesture use predicts later language is less clear. Tomasello (2001) has suggested that the advent of index finger pointing heralds the emergence of new social-cognitive abilities, which allow infants to fully appreciate the function of words as an 'intersubjectively understood linguistic symbol used to direct and share attention with other persons' (Tomasello, 2001: 1120). However, it could simply be that once infants start to use referential gestures, this prompts caregivers to respond by translating their gesture into conventional language at a moment when the infant is jointly attending to both the word and whatever it is denoting (Kishimoto, Shizawa, Yasuda, Hinobayashi, & Minami, 2007). In support of the second view is evidence that infants' first words tend to be names for objects that are labelled using a gesture several months earlier (Iverson & Goldin-Meadow, 2005). Similarly, the reason why non-word repetition performance predicts language learning is ambiguous, with some arguing that the relationship is underpinned by intrinsic phonological short-term memory capacity differences between children (Gathercole, 2006) but others arguing that non-word repetition is simply a measure of the amount and type of linguistic information already stored in the lexicon, which predicts the speed of new vocabulary growth (Jones et al., 2007, see also MacDonald & Christiansen, 2002; Wells et al., 2009). Explicit models that specify how, and why, foundational skills and prior knowledge are causally implicated in individual differences in language growth are essential.

A second unanswered question concerns the origin of the individual differences in the requisite foundational skills. As we argued in section 2.1 above, both our genetic inheritance and our early experiences are likely to affect brain development, and thus influence how efficiently we process, and learn from, linguistic information in infancy and childhood. However, it is notable that some of the correlations between foundational skills and vocabulary development are not present at birth but only emerge during the first year of life. This suggests that children's capacity for learning is not static, or set at birth, but is continually developing, partly as a result of brain maturation (Skeide & Friederici, 2016), but also partly because of the linguistic knowledge they have already accrued in interactions with the communicative environment. For example, Jones and Rowland (2017) have suggested that children who have been exposed to richer linguistic input are able to store, in their lexicon, more lexical and sublexical chunks (sequences of chunked phonemes). This enriched early lexicon enables them both to process familiar words and to learn new words more quickly from the input (because they can use pre-existing chunked knowledge to build lexical representations, rather than building them from scratch, phoneme by phoneme). This results in faster vocabulary growth and a larger vocabulary in later years (Weisleder & Fernald, 2013).

To summarise so far, in this chapter we have argued that individual differences are pervasive across the whole language system, and are thus an essential phenomenon that theories of language acquisition ultimately have to explain. We have described three causal factors implicated in explaining individual differences, which will have to be incorporated into our theories of language acquisition: intrinsic capacity differences in the neurocognitive mechanisms responsible for learning, differences in the child's communicative environment, and differences in the foundational skills upon which language acquisition relies. In the final section we outline a case study which illustrates our approach to individual differences, and highlights the benefits of such an approach in building, testing, and constraining theories of language acquisition.

In this case study, we pit different causal models of individual differences against each other. One possible model (the capacity-limit model) is rooted in capacity-based approaches to cognitive development (e.g. Caplan & Waters, 1999; Just & Carpenter, 1992; Montgomery, Magimairaj, & O'Malley, 2008) and proposes that individual differences in vocabulary acquisition result from intrinsic differences in children's central processing capacity, which causes them to process incoming information more or less efficiently. These differences could emerge, for example, from genetically or environmentally-constrained baseline differences in brain development, which fundamentally shape the brain's capacity for later acquisition. Another (the emergent model) is rooted in experience-based learning theories (e.g. Bates & MacWhinney, 1989; Macdonald, 2013; MacDonald & Christiansen, 2002; Seidenberg & MacDonald, 1999) and proposes that individual differences in lexical processing efficiency are an *emergent* property of children's vocabulary knowledge, in other words, that developments in lexical processing efficiency are driven by changes in long-term knowledge of vocabulary. These differences could be driven by the child's communicative environment, in that richer linguistic environments afford more opportunities for building a lexicon. A third model proposes that lexical processing efficiency reflects a bi-directional causal relationship between the two, with vocabulary improving lexical processing efficiency and lexical processing efficiency improving vocabulary. On this model, the child's capacity for learning can be seen as a continually changing developmental cascade.

3. Case study: Individual differences in children's early lexical processing efficiency.

Here we concentrate on a robust but to-date poorly understood empirical effect: the fact that lexical processing efficiency, as measured by the *looking-while-listening* (LWL) task (Fernald et al., 2006), is significantly associated with children's vocabulary development.

3.1. The LWL task and its relationship to lexical development

The logic behind the LWL task is simple: the child sits facing an eye-tracker or, in a more traditional preferential looking set-up, two screens (Hirsh-Pasek, Cauley, Golinkoff, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996). On each trial, two images appear onscreen (e.g. a car and a fish), and then the child hears audio directing their attention to one of the images (e.g. *Where is the car?*). Several dependent measures can be computed but the most common and intuitive is the average speed with which infants look towards the target on trials when they are initially looking at the distractor image, which we refer to here as *reaction time* (RT). Average RTs have been reported to be associated with individual variability in vocabulary development (Fernald & Marchman, 2012; Fernald, Marchman, & Weisleder, 2013; Fernald et al., 2006; Lany, Giglio, & Oswald, 2018), to predict infants' ability to acquire the meaning of new words (Lany, 2018), and to predict later language and cognitive outcomes (Marchman & Fernald, 2008). Therefore, we concentrate on RTs in the LWL task as our primary measure of lexical processing efficiency.

The demonstration that RTs in the LWL task are associated with language development is well replicated and therefore robust. However, the theoretical nature of the effect is unclear. In their discussion of the effect, Fernald et al. (2006) outlined three possible explanations. Firstly, the effect could reflect a *capacity-based process,* such that efficient lexical processing frees cognitive resources that allow children to process and therefore acquire more words than children with comparatively lower capacities. Such explanations are not uncommon in the psycholinguistic literature investigating the role of working memory in sentence processing in children and adults (Caplan & Waters, 1999; Just & Carpenter, 1992; Montgomery, Magimairaj, & O'Malley, 2008; see Kidd, 2013, for a review), with the assumption being that age-related changes in capacity drive development. We call this the *capacity-limit* model, which is represented in Figure 1a.

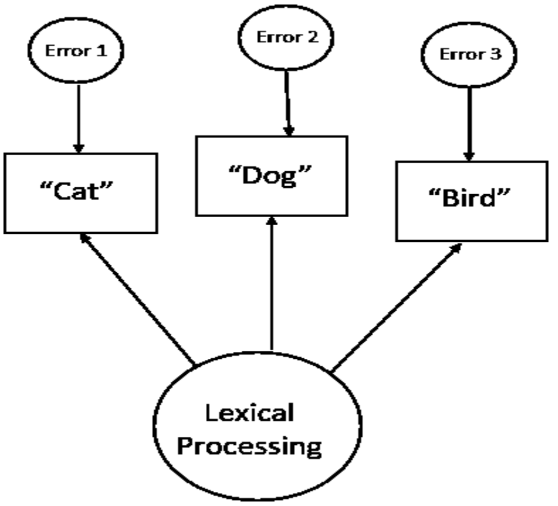

Figure 1 a. Capacity-limit model of lexical processing efficiency.

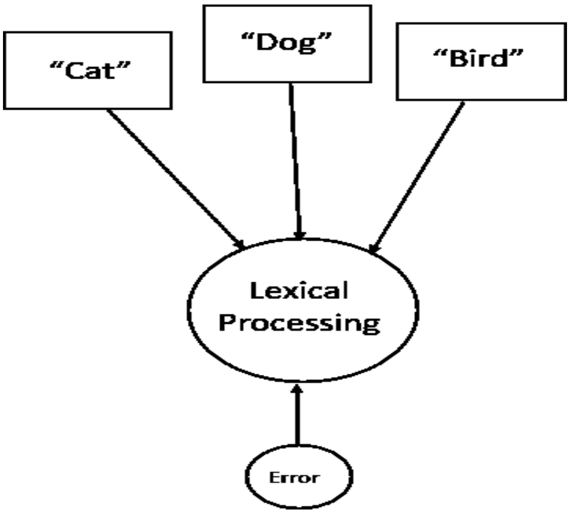

Figure 1 b. Emergent model of lexical processing efficiency.

In Figure 1a, a centralised lexical processing capacity predicts RTs for individual words. These RTs may differ due to extraneous item-level features of the words that are well known in studies of adult lexical access, such as word length, frequency, age of acquisition, and neighbourhood density (Duyck, Desmet, Verbeke, & Brysbaert, 2004). However, importantly, the model predicts that a central processing speed that is independent of individual lexical items is measurable (and indeed statistically *isolable*), contributes to lexical processing speeds, and is predictive of language development.

Secondly, and conversely to the concept of a central processing capacity, children's lexical processing efficiency may be an *emergent* property of their vocabulary knowledge, such that their RTs are simply a reflection of the item-level properties in the lexicon. This alternative model is depicted in Figure 1b, where lexical processing efficiency is defined as the sum of individual RTs on individual words. On this model, development in lexical processing efficiency is driven by changes in long-term knowledge of vocabulary; that is, by developmental changes in the structure of the lexicon. Thus more efficient on-line processing simply reflects more efficient storage and retrieval strategies over development time. The approach is reminiscent of experience-based approaches to language processing, which are mostly based on connectionist approaches to cognition (Bates & MacWhinney, 1989; MacDonald, 2013; MacDonald & Christiansen, 2002; Seidenberg & MacDonald, 1999).

One final possibility, of course, is that lexical processing efficiency reflects a bi-directional causal relationship between the two, with vocabulary improving lexical processing efficiency and lexical processing efficiency improving vocabulary. This third way is consistent with research conducted on the development of processing speed outside the linguistic domain (see Kail & Salthouse, 1994), in which a general assumption is that variance can be both domain-general and domain-specific.

These different theoretical models of lexical processing and its relationship to language development demonstrate the utility and importance of individual differences research in both developing and testing mechanistic models of acquisition. Where correlations such as these are observed, they need to be explained. However, the current evidential base does not yet unambiguously support any of the models described above. There are at least two ways in which we can provide empirical support for one model over the others. Firstly, we can model LWL data using psychometric models that conceptually correspond to the capacity-limit and emergent models, respectively. Secondly, we can test the predictions of each model in a cross-lagged longitudinal study. We describe a study here that does both of these. The data here come from the *Canberra Longitudinal Child Language Project* (CLCL), a longitudinal study of around 130 children acquiring English as a first language in Canberra, Australia (for details see Kidd, Junge, et al., 2018). The study aimed to measure variation in children's on-line language processing skills across development in order to determine how on-line processing interacts with variation in children's input to predict language proficiency. At 18-, 21-, and 24-months the children were tested on a version of the LWL task, and their language development was measured via the MacArthur-Bates Communicative Development Inventory – Words and Sentences (MB-CDI-Words and Sentences, Fenson et al., 2007, for full details see Donnelly & Kidd, 2020).

3.2. Psychometric modelling

The capacity-limit and emergentist accounts of lexical processing efficiency correspond to two classes of models in structural equational modelling, called effects- and causal-indicator models (sometimes called reflective and formative measurement models). The central capacity account corresponds to an effects-indicator model. Effects-indicator models assume that covariation in observed variables is caused by variation in some latent variable (i.e. in our case, a central processing capacity). In the context of the LWL task, mean RTs for individual words would be modelled as shown in Figure 1a, which illustrates the assumptions of an effects-indicator model. The direction of the paths indicates that variation in the latent variable (e.g. a centralised lexical processing capacity) causes variation in the indicator variables (e.g. word-specific reaction times). Because of this, any covariation in the indicator variables is assumed to be due to their shared cause, the latent variable[1]. Moreover, adding or removing indicator variables would not change the interpretation of the latent variable; they are exchangeable indicators of a pre-existing theoretical entity (and thus, lexical processing efficiency should, all things being equal, be roughly equivalent for different words). It is common in psychology to conceptualize constructs in this manner. One example is IQ, where it is traditionally assumed that variation in IQ test scores is caused by *g* (general intelligence; Spearman, 1904) as well as test-specific error. Therefore, increasing *g* would increase the scores on individual IQ tests, but increasing the error components of each individual test

[1] This assumption can be relaxed by allowing their residuals to be correlated.

would not affect *g*. Moreover, adding an additional IQ test to a pre-existing battery would not change the interpretation of *g*.

Treating observed variables as effects indicators is common in psychology, but it is not the only possibility. We could also view the observed variables as causing the latent variable, as is the case, for instance, when a researcher operationalises a social construct like socio-economic status, which is typically derived from numerous indicators such as education level, income, and residence. This conceptualization would lead to a different latent variable model, notably, the emergent model in Figure 1b. This causal-indicator models assumes that RTs for each word combine to create some general processing speed; that is, lexical processing speed is an *emergent* property of knowledge of individual words. In this case, the observed variables are called causal indicators, as opposed to effect indicators. Since they are causally prior to the latent variable, adding or removing causal indicator variables would change the definition of the latent variable (and, thus, different words would result in different processing speeds).

The distinction between causal- and effect-indicator models has been used profitably in other areas of developmental psychology. Notably, Willoughby and colleagues (Willoughby, Blair, & Family Life Project Investigators, 2016) found that a battery of executive functioning tasks administered at ages 3, 4, and 5 was better modelled with causal (i.e. as an emergent process) rather than effects indicators. Moreover, two-year test/re-test reliabilities of this battery differed drastically across the two models. The theoretical implications of such analyses are not to be underestimated. In the cognitive sciences we often take measurements as reflective of stable underlying concepts. For instance, when measuring early vocabulary we assume that the MB-CDI and naturalistic samples of speech reflect the same underlying knowledge base, or when measuring Theory of Mind we assume that the unexpected location and the appearance-reality tasks reflect false belief reasoning (Wellman, 2014). However, it is equally possible that the concepts are emergent properties of the measures themselves, and psychometric modelling of the data allows us to test for this possibility.

In Donnelly and Kidd (2020), we did just that for lexical processing efficiency in the CLCL cohort at 18 months. At this time point, the children were tested on eight target words, six times each. The data were collected using a Tobii 60XL eyetracker, sampling at 60Hz. For each trial the children saw two pictures (the target, and a distractor, which was taken from the seven remaining words used in the experiment) appear on screen. After 2000ms, an audio file, recorded by a female native speaker of Australian English, directed the child to the target image. The entire experiment lasted approximately 7 minutes. We collected enough valid trials from 95 children, whose data were fit using structural equation models representing the capacity-limit and emergentist models, respectively. The comparative fit of the two models was then compared to determine which best explained the data.

Interestingly, *both* model types resulted in excellent model fit, and a statistical comparison of the two models did not provide strong support for one model over the other. The results stayed largely the same when we modelled a different DV – proportion of time looking to the target. Furthermore, preferring one model over the other would not have led to different inferences about lexical processing efficiency. For instance, modelling lexical processing efficiency as either a capacity-based or emergent skill resulted in very similar relationships between 18-month lexical processing and lexical processing and vocabulary at 21 months. Thus, given these initial analyses, we did not have enough evidence to distinguish between either model of lexical processing efficiency. A larger sample size may provide a more rigorous test between these two approaches, but the overall fit statistics suggest that either account is empirically plausible. Moreover, the similar relationships between lexical processing efficiency at 18 months and the 21 month outcomes across the two models suggest that researchers' choice of conceptualization, while theoretically significant, should not lead to different conclusions about the relationship between lexical processing efficiency and later outcomes.

3.3. Longitudinal analyses

Our next analyses investigated the relationship between lexical processing efficiency and vocabulary, as measured by the MB-CDI, at 18-, 21-, and 24-months of age. Both concepts were measured at each time point in a cross-lagged design, such that we could determine the relationship between lexical processing efficiency and vocabulary longitudinally while controlling for the effects of both variables at earlier points in development. While our initial intention was to estimate a random-intercepts cross-lagged panel model, this was not possible because of the strong non-linear relationship between vocabulary at consecutive time points In order to test whether performance on the LWL predicts vocabulary over and above prior vocabulary, this non-linear relationship must be correctly modelled. We, therefore, conducted separate regressions on vocabulary at 21 and 24 months to identify a suitable relationship between vocabulary at consecutive time points. Vocabulary at 21 months was best modelled using linear and quadratic terms for 18 month vocabulary, and vocabulary at 24 months was best modelled using linear, quadratic and cubic terms for 21 month vocabulary. Once these relationships had been appropriately modelled, we added LWL variables to the models to see if they explained any additional variation. For 21 month vocabulary, LWL RTs did not significantly predict vocabulary, but LWL proportions did. For 24 month vocabulary, neither LWL measure was significant.

We then examined whether vocabulary predicted LWL over and above prior LWL. RTs at 21 months were significantly predicted by vocabulary at 18 months, over and above 18-month RTs. Proportion of looks to target at 21 months was marginally significantly predicted by prior vocabulary over and above prior LWL. Neither 24-month LWL measure was significantly predicted by 21-month vocabulary over and above prior LWL. These relationships are summarised

Overall the results hint at a bidirectional relationship between vocabulary and lexical processing efficiency, at least from 18 to 21 months. There was moderate evidence for an effect of lexical processing efficiency at 18 months on vocabulary at 21 months, as the effect was significant when lexical processing efficiency was measured as a proportion though not as an RT. There was even stronger evidence for an effect of vocabulary at 18 months on lexical processing efficiency at 21 months, as the effect was significant for RTs and marginally significant for proportions. But the effects of both disappeared between 21 and 24 months. An important question, then, is why the relationship would exist between 18 and 21 months and disappear between 21 and 24 months?

The moderate evidence that lexical processing efficiency was related to growth in vocabulary is consistent with work indicating that children with better lexical processing efficiency perform better in novel word learning tasks than children with lower lexical processing efficiency at 17 months (Lany, 2018). However, why did this effect disappear between 21 and 24 months? Lany (2018) observed that lexical processing efficiency at 30 months predicted word-learning in challenging contexts. One possibility is that the extremely high correlation between vocabulary at 21 and 24 months ($r = .89$) rendered such an effect impossible to observe.

A similar pattern was observed in the models examining the effect of vocabulary on LPE. There was strong evidence that children with larger vocabularies at 18 months exhibited more LPE growth between 18 and 21 months than did children with smaller vocabularies, an effect that was significant for RTs and marginally significant for proportions. At first blush this account seems to provide evidence for the emergentist account of lexical processing efficiency, but why the relationship disappears between 21 and 24 months is unclear. One possible explanation for the developmental nature of the effect concerns the changing structure of the infant lexicon. Until this point, research in this space has been agnostic about how words are represented and accessed. However, as children rapidly acquire words, the structure of their lexicon changes, and this is likely to have a significant impact on lexical access. Words are stored in the lexicon according to the phonological and semantic similarity of their neighbourhood structure. In adults, semantic and phonological neighbourhood effects are commonly observed and are explained by interactive-activation and competition models of lexical access (for review see Chen & Mirman, 2012). Importantly, neighbourhood structure can have both facilitative and inhibitory effects on lexical access depending on neighbourhood density and the particular linguistic task (e.g., production versus comprehension). For example, phonological neighbourhood density inhibits word recognition but facilitates word production; the number of distant semantic neighbours (e.g., *bread, cereal*) facilitates word recognition while the number of near semantic neighbours (e.g., *bread, butter*) inhibits it. Chen and Mirman (2012) presented a single process model of lexical processing that incorporates activation and inhibitory processes: strongly active neighbours inhibit access to the target representations whereas weakly active neighbours facilitate access.

As infants learn new words, the structure of the lexicon changes, leading to more densely populated phonological and semantic neighbourhoods (for a review see Wojcik, 2018). However, the current state of research suggests that these two variables may come online at different points in development. Rämä, Sirri, and Serres (2013) observed semantic priming in 24-month-old infants as indexed by an N400 in an experiment using event-related potentials, and observed a similar effect in 18-months-olds who had comparatively high vocabularies for their age. Thus it appears that semantic neighbourhood density begins to have an effect on lexical access before a child’s second birthday. In contrast, studies investigating the effect of phonological neighbourhood density on infant lexical access suggest that the lexicon may begin to become interconnected some months later, such that phonological neighbourhood density effects are reliably observed at 24 months (Mani & Plunkett, 2010, 2011). This apparent decalage between the emergence of semantic and phonological structure in the lexicon could explain our observed developmental relationship between lexical proficiency and vocabulary development. If semantic relatedness begins to influence lexical processing at 18 months, this could explain the early relationship we found between vocabulary size and processing efficiency. Specifically, it could be that high-vocabulary children accessed target items more quickly because their lexicons are more semantically interconnected in such a way that greater semantic density leads to faster RTs on the LWL task.

Relatedly, the later onset of phonological interconnectedness may be responsible for the fact that we did not observe a similar effect of 21-month vocabulary predicting 24-month lexical processing. Mani and Plunkett (2011) have shown that greater phonological cohort density can have an inhibitory effect on lexical access. Thus, our data could be explained by the countervailing influence of semantic and phonological network structure on lexical processing: early in development, vocabulary growth could facilitate lexical processing by increasing semantic neighbourhood density, but later growth may be unrelated due to the emerging inhibitory influence of phonological neighbourhood density. This explanation crucially relies on whether there is in fact a developmental decalage between the emergence of semantic and phonological structure in the infant lexicon. It does seem to be the case that infants preferentially learn from phonologically and semantically dense neighbourhoods (Carlson, Sonderegger, & Bane, 2014; Hills, Maouene, Maouene, Sheya, & Smith, 2009), but the question of whether semantic neighbourhood density emerges first awaits future research.

4. Future prospects

Our explanation of the changing longitudinal relationship between lexical processing efficiency and vocabulary clearly implicates long-term lexical and sub-lexical knowledge as a driver of processing, and, on face value, is more closely aligned with an emergent conception of processing in comparison to a capacity-limit account. That is, we have explained lexical processing efficiency as emerging from item-level effects in the lexicon. Changes in long-term knowledge of vocabulary are likely driven in part by the richness of the child's communicative environment, in that richer linguistic environments afford more opportunities for building a lexicon. Thus, our case study suggests that a richer input affects language acquisition not only directly, by providing the children with more opportunities for learning a greater diversity of words, but also indirectly, by increasing processing capacity, and thus speeding up the learning of new words. We also note that our data show evidence of a clear developmental cascade, since the relationship between lexical processing speed and vocabulary changes over developmental time, perhaps as a result of the changing nature of the structure of the lexicon.

However, while our data likely rule out an explanation of lexical processing that is completely independent of lexical knowledge, it does not rule out the possibility of intrinsic individual differences in the efficiency of a centralised processing capacity *in addition to* item-level knowledge. For instance, it is possible that genetically or early environmentally-driven individual differences in the neurological development such as myelination or the number of neural connections devoted to linguistic processing lead to faster processing over and above long-term knowledge of language. Such possibilities await future research.

Finally, we return to Elena Lieven’s work. There are two limitations of the current body of work on individual difference, both noted by Elena Lieven years ago. First, most of the work on individual differences has been conducted with children learning a limited number of languages; mainly English, though with some work in other WEIRD countries (Western Educated, Industrialised Rich and Democratic countries; e.g. on Spanish, Hurtado, Marchman, & Fernald, 2007, or Danish, Bleses et al., 2008). This limits the scope and generalisability of our theory building (Lieven, 1994). For example, if the lexicon plays a central role in determining the speed of word processing and learning, as suggested above, then we might predict cross-linguistic differences based on the semantic and phonological shape of the lexicon in different languages (see Frank et al., n.d. for some initial findings). Further work on a range of languages is required to test this hypothesis. Second, much of the work cited in this chapter has almost exclusively measured children’s language development using parental report checklists such as the MacArthur CDI (Fenson et al., 2007). However, as Pine, Lieven, & Rowland (1996) showed over 20 years ago, while we can expect checklists to provide good approximations of children’s language, making them good instruments to sample individual differences, we cannot expect them to provide good estimates of the relative distribution of different lexical classes in a child’s lexicon. Most notably, they tend to overestimate the proportion of nouns in a child’s vocabulary. This is not a problem per se for research using the LWL paradigm. However, if our suggestion is that the composition of a child’s lexicon has a direct bearing on the speed of future language growth, then future studies need to estimate the structure of the individual child’s lexicon with a high level of accuracy, which may require additional complementary measures (e.g. lab-based elicitation or comprehension studies to provide detailed information about knowledge in particular semantic domains) and additional types of analyses (e.g. Borovsky, Ellis, Evans, & Elman, 2016's analysis of the semantic structure of children's vocabularies) .

References

Abrahams, B. S., Tentler, D., Perederiy, J. V., Oldham, M. C., Coppola, G., & Geschwind, D. H. (2007). Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proceedings of the National Academy of Sciences*, *104*(45), 17849–17854.

Alarcón, M., Abrahams, B. S., Stone, J. L., Duvall, J. A., Perederiy, J. V., Bomar, J. M., … Geschwind, D. H. (2008). Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *The American Journal of Human Genetics*, *82*(1), 150–159.

Ambridge, B. (2018). Against stored abstractions: A radical exemplar model of language acquisition. Retrieved from https://ling.auf.net/lingbuzz/004131

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*(02), 239–273.

Bates, E., Bretheron, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. New York: Cambridge University Press.

Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. Macwhinney (Eds.), *The Handbook of Child Language* (pp. 96–151). Oxford: Basil Blackwell.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes*, *12*(5/6), 507–584.

Bates, E., & Macwhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The Crosslinguistic Study of Sentence Processing* (pp.

3–76). Cambridge: Cambridge University Press.

Beals, D. E. (1997). Sources of support for learning words in conversation: evidence from mealtimes. *Journal of Child Language*, *24*(3), 673–694.

Bishop, D. V. M. (2003). *Test for Reception of Grammar: TROG-2 version 2*. London: Pearson Assessment.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). The Danish Communicative Developmental Inventories: validity and main developmental trends. *Journal of Child Language*, *35*(03), 651–669.

Bloom, L., Lightbown, P., Hood, L., Bowerman, M., Maratsos, M., & Maratsos, M. P. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, *40*(2).

Bloomfield, F. H. (2011). Epigenetic modifications may play a role in the developmental consequences of early life events. *Journal of Neurodevelopmental Disorders*, *3*(4), 348–55.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123–172). Dordrecht: Reidel.

Bornstein, M. H., Haynes, M. O., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of Child Language*, *25*(2), 367–393.

Bornstein, M. H., & Putnick, D. L. (2012). Stability of language in childhood: A multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology*, *48*(2), 477–491.

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure invocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893–1908.

Brito, N. H., Fifer, W. P., Myers, M. M., Elliott, A. J., & Noble, K. G. (2016). Associations among family socioeconomic status, EEG power at birth, and cognitive skills during infancy. *Developmental Cognitive Neuroscience*, *19*, 144–151.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77–94.

Carlson, M. T., Sonderegger, M., & Bane, M. (2014). How children explore the phonological network in child-directed speech: A survival analysis of children's first word productions. *Journal of Memory and Language*, *75*, 159–180.

Cartmill, E., Armstrong, B. F., Gleitman, L. R., Goldin-Meadows, S., Medinac, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, *110*(28), 11278–83.

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*(2), 417–430.

Chonchaiya, W., Tardif, T., Mai, X., Xu, L., Li, M., Kaciroti, N., … Lozoff, B. (2013). Developmental trends in auditory processing can provide early predictions of language acquisition in young infants. *Developmental Science*, *16*(2), 159–172.

Choudhury, N., Leppanen, P. H. T., Leevers, H. J., & Benasich, A. A. (2007). Infant information processing and family history of specific language impairment: Converging evidence for RAP deficits from two paradigms. *Developmental Science*, *10*(2), 213–236.

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352–366.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for Specific Language Impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741–748.

Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: A study of young language-learning children. *International Journal of Language and Communication Disorders*, *38*, 251–263.

Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America*, *129*(5), 3271–80.

Cristià, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development*, *85*(4), 1330–1345.

Dąbrowska, E., & Street, J. A. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*(6), 604–615.

Demir-Vegter, S., Aarts, R., & Kurvers, J. (2014). Lexical richness in maternal input and vocabulary development of Turkish preschoolers in the Netherlands. *Journal of Psycholinguistic Research*, *43*(2), 149–65.

Donnelly, S., & Kidd, E. (2020). Individual differences in lexical processing efficiency and vocabulary in toddlers: A longitudinal investigation. *Journal of Experimental Child Psychology*.

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *36*(3), 488–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15641437

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, *43*(4), 865–

78.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*(2), 321–35.

Falcaro, M., Pickles, A., Newbury, D. F., Addis, L., Banfield, E., Fisher, S. E., Monaco, A. P., Simkin, Z., Conti-Ramsden, G. & The SLI Consortium (2008). Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment. *Genes, Brain and Behavior*, *7*(4), 393–402.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual.* (2nd ed.). Baltimore, MD: Brookes Publishing Co.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203–222.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116.

Festman, J., Rodriguez-Fornells, A., & Münte, T. F. (2010). Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behavioral and Brain Functions, 6*(1), 5.

Frank, M. C., Braginsky, M., Marchman, V., & Yurovsky, D. (n.d.). *Variability and Consistency in Early Language Learning*. Cambridge, MA: MIT Press. Retrieved from https://langcog.github.io/wordbank-book/index.html

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513–543.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*(5), 887–898.

Harris, J. R. (1998). *The nuture assumption: Why children turn out the way they do*. New York: Free Press.

Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, *28*(6), 1096–1105.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes Publishing Co.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729–39.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., … Suma, K. (2015). The contribution of early communication quality to low-income children’s language success. *Psychological Science*, *26*(7), 1071–1083.

Hirsh-Pasek, K., Cauley, K. M., Golinkoff, R. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(1), 23–45.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Hoareau, M., Yeung, H. H., & Nazzi, T. (2019). Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities. *Developmental Science*.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development VO - 73*, (2), 418.

Hoff, E., Quinn, J. M., & Giguere, D. (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*, *21*(2), e12536.

Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language*, *44*(1), 63–86.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, *33*, 227–249.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31–F39.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*, 343–365.

Hyams, N. M. (1986). *Language acquisition and the theory of parameters*. Dordrecht: D. Reidel.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367–371.

Johnson, W., Turkheimer, E., Gottesman, I. I., & Bouchard, T. J. J. (2010). Beyond heritability: Twin studies in behavioral research. *Current Directions in Psychological Science*, *18*(4), 217–220.

Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, *10*(6), 853–873.

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology*, *98*, 1–21.

Junge, C., Cutler, A., & Hagoort, P. (2012). Electrophysiological evidence of early word learning. *Neuropsychologia*, *50*(14), 3702–3712.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–49.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*(2–3), 199–225.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398.

Khanna, M. M., & Boland, J. E. (2010). Children’s use of language context in lexical ambiguity resolution. *Quarterly Journal of Experimental Psychology*, *63*(1), 160–193.

Kidd, E. (2013). The role of verbal working memory in children’s sentence comprehension. *Topics in Language Disorders*, *33*(3), 208–223.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169.

Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual differences in infant speech segmentation: Achieving the lexical shift. *Infancy*, *23*(6), 770–794.

Kishimoto, T., Shizawa, Y., Yasuda, J., Hinobayashi, T., & Minami, T. (2007). Do pointing gestures by infants provoke comments from adults? *Infant Behavior and Development*, *30*(4), 562–567.

Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V. M., & Plomin, R. (2005). Genetic influences in different aspects of language development: The etiology of language skills in 4.5-year-old twins. *Child Development*, *76*(3), 632–651.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000.

Lany, J. (2018). Lexical-processing efficiency leverages novel word learning in infants and toddlers. *Developmental Science*, *21*(3), e12569.

Lany, J., Giglio, M., & Oswald, M. (2018). Infants' lexical processing efficiency is related to vocabulary size by one year of age. *Infancy*, *23*(3), 342–366.

Law, J., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health and employment outcomes. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1401–1416.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language* *acquisition* (pp. 56–73). Cambridge, UK: Cambridge University Press.

Lieven, E. V. M. (2016). Usage-based approaches to language development: Where do we go from here? *Language and Cognition*, *8*, 346– 368.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1–F10.

Macdonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35-54; discussion 55-74.

Mani, N., & Plunkett, K. (2010). In the infant's mind's ear. *Psychological Science*, *21*(7), 908–913.

Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*(2), 196–206.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9-16.

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, *22*(03), 491–495.

Matthews, D., Behne, T., Lieven, E., & Tomasello, M. (2012). Origins of the human pointing gesture: A training study. *Developmental Science*, *15*(6), 817–829.

McGillion, M. L., Herbert, J. S., Pine, J. M., Keren-Portnoy, T., Vihman, M. M., & Matthews, D. E. (2013). Supporting early vocabulary development: What sort of responsiveness matters. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 240–248.

McGillion, M. L., Herbert, J. S., Pine, J., Vihman, M. M., Depaolis, R., Keren-Portnoy, T., & Matthews, D. (2016). What paves the way to conventional language? The predictive value of babble, pointing and SES. *Child Development*, *88*(1), 156–166.

Montgomery, J. W., Magimairaj, B. M., & O'Malley, M. H. (2008). Role of working memory in typically developing children's complex sentence comprehension. *Journal of Psycholinguistic Research*, *37*(5), 331–354.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, *38*(1–2).

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience*, *29*(2), 186–217.

Nozari, N., Trueswell, J. C., & Thompson-Schill, S. L. (2016). The interplay of local attraction, context and domain-general cognitive control in activation and suppression of semantic distractors during sentence comprehension. *Psychonomic Bulletin and Review*, *23*(6), 1942–1953.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*(4), 763–82.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41–58.

Pine, J. M., & Lieven, E. V. M. (1993). Reanalysing rote-learned phrases: individual differences in the transition to multi-word speech. *Journal of Child Language*, *20*(03), 551–571.

Pine, J. M., Lieven, E. V. M., & Rowland, C. (1996). Observational and checklist measures of vocabulary composition: what do they mean? *Journal of Child Language*, *23*(3), 573–

590.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early grammars of English*. Oxford: Blackwell.

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18- and 24-month old children. *Brain and Language*, *125*(1), 1–10.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*(4), 850–63.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11 -month-old American infants. *Developmental Science*, *8*(2), 162–72.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710.

Roth, C., Magnus, P., Schjølberg, S., Stoltenberg, C., Surén, P., McKeague, I. W., … Susser, E. (2011). Folic acid supplements in pregnancy and severe language delay in children. *JAMA*, *306*(14), 1566.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23*(4), 569–588.

Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network.

Nature Reviews Neuroscience, *17*(5), 323–332.

Spearman, C. (1904). “General Intelligence,” objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–293.

Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of Specific Language Impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research*, *49*(2), 219–236.

Street, J. A., & Dabrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*, *12*(8), 2080–2094.

Street, J. A., & Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers’ processing of the English passive. *Applied Psycholinguistics*, *35*(1), 97–118.

Tomasello, M. (2001). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*, 61–82.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084.

van der Lely, H. K. J., & Pinker, S. (2014). The biological basis of language: insight from developmental grammatical impairments. *Trends in Cognitive Sciences*, *18*(11), 586–595.

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., … Fisher, S. E. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, *359*(22), 2337–2345.

Vuong, L. C., & Martin, R. C. (2014). Domain-specific executive control and the revision of misinterpretations in sentence comprehension. *Language, Cognition and Neuroscience*, *29*(3), 312–325.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265–279.

Wellman, H. (2014). *Making minds: How theory of mind develops.* Oxford: Oxford University Press.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271.

Whitehouse, A. J. O., Bishop, D. V. M., Ang, Q. W., Pennell, C. E., & Fisher, S. E. (2011). CNTNAP2 variants affect early language development in the general population. *Genes, Brain and Behavior*, *10*-(4), 451–456.

Willoughby, M. T., Blair, C. B., & Family Life Project Investigators, the F. L. P. (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment*, *28*(3), 319–30.

Wojcik, E. H. (2018). The development of lexical-semantic networks in infants and toddlers. *Child Development Perspectives*, *12*(1), 34–38.

Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology*, *141*, 187–209.

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience*

& Biobehavioral Reviews, *81*, 103–119.