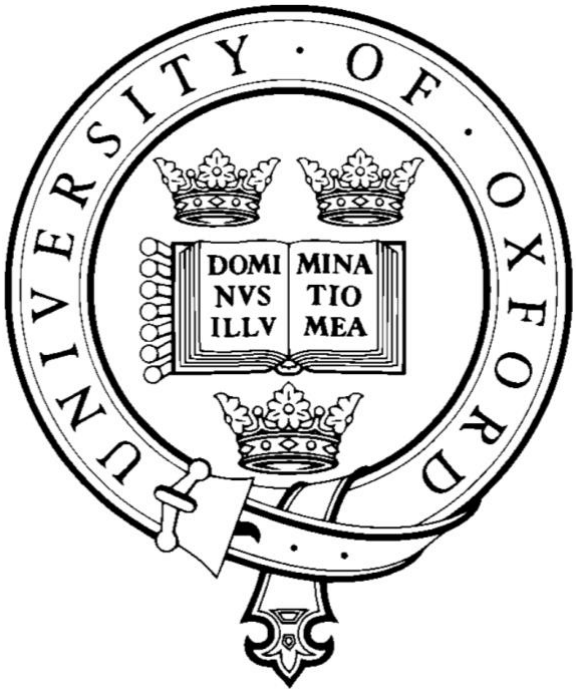
Emotion Language in Context: Investigating Emotion Words in Children's Written Language and Word Learning



Yuzhen Dong

St John's College

University of Oxford

Supervised by:

Prof. Kate Nation and Prof. Robert Hepach

A thesis submitted to the Department of Experimental Psychology, University of Oxford, in partial fulfilment for the degree of Doctor of Philosophy

January 2025

Front Matter

Declaration

I declare that this thesis is entirely my own work, and except where otherwise stated, describes my own research. In line with open science practice, all the materials, data, and scripts associated with this thesis have been made available on the Open Science Frameworks (OSF).

Chapter 2: http://tiny.cc/rainy_thesis_chapter2

Chapter 3: http://tiny.cc/rainy_thesis_chapter3

Chapters 4 and 5: http://tiny.cc/rainy_thesis_chapter4and5

Chapter 6: http://tiny.cc/rainy_thesis_chapter6

Publications and submitted preprints arising from the thesis

Some of the work reported in this thesis has been published or submitted for publication. For those chapters, I maintain usage of 'we' and 'our' reflecting the multiple authorship of the papers, with my co-authors providing constructive feedback (Chapters 2-5). In my General Introduction, Chapter 6, and General Discussion chapters, I revert to the first person, reflecting sole authorship of these sections.

Chapter 2 has been uploaded to OSF as a preprint and submitted to *First Language*:

Dong, Y., & Nation, K. (2024). *Charting the frequency and diversity of emotion words in children's language: Written language matters*. Submitted to First Language (19th September 2024); Major revision decision (29th October 2024); pre-print available at: OSF. https://doi.org/10.31219/osf.io/zj5k7

Chapter 3 describes work published in *Cognitive Science*:

Dong, Y., Hsiao, Y., Dawson, N., Banerji, N., & Nation, K. (2024). The Emotional Content of Children's Writing: A Data-Driven Approach. *Cognitive Science*, *48*(3), e13423. https://doi.org/10.1111/cogs.13423

Chapter 4 describes work published (in press) in *Quarterly Journal of Experimental Psychology*:

Dong, Y., Mak, M. H., Hepach, R., & Nation, K. (2024). EXPRESS: Learning New Words via Reading: The Influence of Emotional Narrative Context on Learning Novel Adjectives. *Quarterly Journal of Experimental Psychology*, 0(0). https://doi.org/10.1177/17470218241308221

Chapters 5 and 6 will be developed into journal articles after thesis submission.

Acknowledgement

This DPhil journey would not have been possible without the support, guidance, and encouragement of many people along the way. I feel extremely fortunate and grateful to have met so many supportive mentors and dear friends.

Thank you, Kate, for being the most incredible and inspirational supervisor. You were there for every "first" step of my academic journey, my first talk, conference, and publication just to name a few. Your academic guidance and generous encouragements have driven my development as a psychologist and a scientific thinker. It has been a great honour and privilege to learn and grow under your mentorship over the past three years.

To Robert, thank you for being a wonderful co-supervisor and offering fresh perspectives on my work. Our discussions on the emotional aspects of my research have always been inspiring. A heartfelt thank you also to Liz, my college advisor, for providing thoughtful advice on both my research and career, and for inviting me to SCR lunches.

To the ReadOxford team, thank you all for creating such a friendly and encouraging environment to do fun research. Special thanks to Nicky, Yaling, and Matthew for being brilliant mentors and collaborators, always generous with your wisdom and expertise. To Mohen, Sean, Nicole, Marlene, Jessie, Tonia, and Clemens, thank you for being my lovely academic siblings. Your camaraderie and assistance make coming to AW something to look forward to.

I'm also grateful to the wider Psychology community. To members of the Nation-Wonnacott joint lab group and the Developmental Hour group, thank you for listening to my half-baked designs and practice talks. My appreciation also goes to Dan, Martin, and Vinca for their meticulous administrative support, and to the six undergraduate Research Assistants who helped with school recruitment for my final project. Completing data collection for 120 children within a month felt nothing short of a miracle, and it could not have happened without your dedication.

Thank you to St John's College for supporting my academic journey through generous travel awards and grants, and to the wonderful community of people I meet here, for being the cosy place which I now call my second home. Thank you to China Oxford Scholarship Fund for providing the financial support and connecting me to a community of great academics.

To Yuanyuan, Shuyi, and Siyu, thank you for being my first and best friends at Oxford, and for making my DPhil experience more colourful.

Finally, to my family, my deepest gratitude goes to my parents. Thank you for creating the best environment for me to thrive, for always believing in me, and for nurturing me into a kind and positive person. To my fiancé Tony, thank you for your constant presence, countless happy memories, and for the love and commitment that keep me grounded. I love you.

Abstract

The overarching aim of my thesis is to examine emotion language in children's reading and writing, and the learning of new emotion words by adults and children. Understanding emotion words is crucial for emotional concept development, but few studies have considered emotion words beyond early childhood, or in the context of written language. This is an important gap as children experience language through reading emotional narratives, and this brings opportunities to learn new words and abstract concepts. Little is known about the potential influence that the emotional valence of the text might have on word learning. The thesis addresses these gaps in the literature through a combination of data-driven corpus analyses and experimental work.

Chapter 1 introduces the broad topics of language and emotion, and the scope of the thesis. Three sets of data-driven corpus analyses (Chapters 2,3, and 5) examine the frequency and distribution of emotion words in texts written by and for children, as well as the emotional valence associated with these words. Chapter 2 reports cross-corpus comparisons showing that language written by and for older children contained a greater diversity of emotion words than those written by and for younger children. After controlling for age, even books targeted at pre-schoolers for shared reading contained more unique emotion words than both child-directed speech and television language. In Chapter 3, sentiment analyses across a large sample of children's writing revealed that positive sentiments in children's writing decreased with age. The stories were analysed further in Chapter 5, where there was a positive correlation between word valence and context valence, a pattern that was consistent across ages.

To complement the corpus approach, this thesis also includes two word-learning experiments, one with adults (Chapter 4) and one with children (Chapter 5). Participants read novel adjectives in short narratives that had either neutral, negative, or positive valence. Adults and children were able to infer the valence of the novel words from reading, and those words read in either positive or negative contexts were learned better than those words read in the neutral context. Chapter 6 integrates multiple approaches to explore how children's experience with reading and written language is related

to the emotional content of the language they produce. Preliminary analyses revealed that reading experience was a significant predictor of the lexical diversity of their overall narrative production, but not for the production of emotion words specifically, despite the positive correlation between reading experience and emotion word production. Chapter 7 closes with a general discussion that offers reflections and suggestions for future research.

Together, my thesis combines both data-driven analyses and experimental work to better understand the emotional content in children's written language and how emotional contexts influence word learning. The research addresses a knowledge gap in written language and emotion, which holds implications for language and literacy development, emotional well-being, and educational practice.

Table of Content

List of Figures

List of Tables

Chapter 1. General Introduction

Imagine a child excitedly describing the joy of a birthday party to their parents. They might jump, wave their hands, and speak in a high-pitched voice to express how happy they feel. Their wide eyes and broad smiles would further reveal their emotions. Now, picture the same child attempting to convey this joy to a pen friend via email. Without the aid of vocal tone, gestures, or facial expressions, the child faces a more challenging task: relying solely on language to express their feelings. In writing about this event, they may need to craft precise and nuanced language to convey the same joy.

The emotion language a child chooses to use might offer a glimpse into their emotional world. The ability to use specific emotion words allows children to better recognise and regulate their emotions, and navigate social interactions (e.g., Streubel et al., 2020). Early in life, infants acquire emotion words mostly from hearing caregivers' speech (e.g., Ogren & Sandhofer, 2021). As literacy develops in mid-childhood, reading and writing opens new opportunities to learn and engage with emotion words through written language (Nation et al., 2022). What kinds of emotion words are found in books, and how do they differ from those in everyday conversation? How are emotion words used by children, and how does emotional context influence language learning? This thesis explores these questions about emotion language from multiple angles, by examining how children use and acquire emotion words and how their language environment shapes this process. The research aims to deepen our understanding of the relationship between language and emotion and its broader implications for psychological well-being. This chapter offers a broad overview that outlines some critical background and sets out the scope of the thesis. A more detailed review and evaluation of relevant literature and methods will be provided in each of the chapters that follow.

Background: language and emotion

Language development is a dynamic process that begins in early infancy and continues throughout life. As the building block for language, words serve not only as labels for objects and actions but also help infants learn context-dependent, conceptual categories (Gelman & Roberts, 2017). While

concrete categories such as objects with perceptual similarities can be learned through statistical regularities and capacities such as cross-situation learning (e.g., Romberg & Saffran, 2010; Saffran et al., 1996; Saffran & Kirkham, 2018; Seidenberg & MacDonald, 2018), categories that are more abstract such as emotion and mental states are more challenging to grasp. In these cases, words act as "conceptual glue" that binds diverse sensory experiences into coherent categories as caregivers label objects, actions, or events. These interactions help infants organise both concrete and abstract concepts, providing scaffolds for their cognitive development and understanding the complexities of the world (Gelman & Roberts, 2017; Xu, 2007; Xu & Tenenbaum, 2007).

Similar to language, development of emotion also begins at birth (for review, see Ruba et al., 2022). Emotions guide how we respond to our environment and influence our behaviour, decision-making, and social interactions (LoBue et al., 2019). Greater proficiency in perceiving emotions in others, alongside a deeper understanding of emotional concepts, is associated with increased subjective well-being and enhanced emotion regulation (Kashdan et al., 2015). These emotional abilities are also linked to greater prosocial behaviour (Eggum et al., 2011), higher educational attainment, and improved productivity and workplace relationships (Brackett et al., 2012; Hagelskamp et al., 2013).

The nature of emotions has been debated across various theoretical perspectives. The categorical view posits that emotions can be grouped into discrete, basic, and universal categories, such as anger, sadness, joy, fear, surprise, and disgust (Ekman, 1992). These categories are believed to be biologically hardwired and universally recognisable across cultures. The dimensional view, on the other hand, states that emotions exist along continuous dimensions, such as valence (positive or negative affect) and arousal (high or low intensity) (Russell, 1980). From this perspective, emotions are not viewed as discrete entities but rather as points within a multidimensional space, which highlights the fluidity of emotional experiences. Beyond these two approaches, psychological constructionist views propose that emotions are not biologically predefined but instead arise from the combination of multiple cognitive and affective processes, constructed by the mind to fit specific situations (Cunningham et al., 2013; Russell, 2003). For example, the experience of fear may involve

different physiological and behavioural responses depending on whether the threat is social or physical (Lindquist & Barrett, 2008; LoBue & Adolph, 2019). A related theory, the theory of constructed emotions states that emotion categories are abstract, goal-based groupings of emotional instances, with features that are dynamically constructed by the brain (Barrett, 2006, 2017). According to this view, language plays an important role in shaping how we conceptualise and experience emotions.

In the following sections, I review how language and emotion supports the development of each other.

Language in emotional development

In line with the constructionist view, Hoemann et al. (2019) proposed that language plays a critical role in the acquisition of emotional knowledge and is central to developing emotion understanding (Barrett, 2004, 2017; Hoemann et al., 2019; Lindquist, 2017; Shablack & Lindquist, 2019). According to this view, emotional development is not simply the maturation of innate emotional responses but a process of constructing emotion concepts through social interactions that are contextualised by language (Lutz, 1982; Ratner, 1989). Language allows children to connect diverse affective, physical, and perceptual experiences with specific emotion categories, facilitating the development of complex emotional knowledge. For example, hearing the word “angry” in various contexts where they experience frustration or obstruction helps children categorise these diverse experiences under the label of “anger”. This process of labelling with emotion words extends beyond recognising facial expressions or physiological responses; it allows children to understand the functional role that emotions play in different situations, even when those situations may not have clear perceptual similarities. In this way, language not only labels emotions but also helps children make sense of their varied and often ambiguous emotional experiences (Hoemann et al., 2019; Lindquist, 2017; Lindquist et al., 2015; Shablack & Lindquist, 2019).

Research supports the idea that emotion words shape emotional perception and concept formation. Doyle and Lindquist (2018) demonstrated that pairing facial expressions with emotion words

influences how these expressions are later remembered and perceived. Similarly, Gendron et al. (2012) found that when the meaning of an emotion word (e.g., “anger”) was temporarily suppressed by repeating it 30 times, participants were less accurate at categorising related emotional facial expressions. This diminished performance highlights how emotion words support the conceptual understanding of emotions, even when tasks do not explicitly involve language. The temporary inaccessibility of an emotion word disrupts the priming of emotional cues, showing how tightly linked language is with emotional perception.

The relationship between language and emotional development is further demonstrated by evidence showing that children’s acquisition of emotion words is associated with their ability to understand and categorise emotions. For example, Russell and Widen (2002) found that children aged 3 to 4 years were better able to categorise facial expressions when emotion labels were provided, compared to when they were not. Similarly, Widen and Russell (2008) found that children who knew more emotion words were more accurate in sorting facial expressions. This suggests that as children expand their emotion vocabulary, their emotional concepts become more nuanced.

There is also evidence that both the breadth and depth of emotion vocabulary contribute to emotional development. Streubel et al. (2020) found that the size of preschool-aged children’s emotion-specific vocabulary was linked to their knowledge of emotion regulation strategies. Among primary school-aged children, the depth of their emotion vocabulary predicted individual differences in emotion understanding. As children grow older, their use of emotion words becomes more specific and adult-like (Grosse et al., 2021). Additionally, in line with the view that language is critical to the construction of emotion, there is also an association between language competence and emotion understanding in pre-schoolers (e.g., Widen & Russell, 2003, 2008) and in older children (e.g., Beck et al., 2012; Griffiths et al., 2020). Thus, there is evidence that language supports emotional development by helping children categorise, understand, and regulate their emotional experiences.

Emotion in language development

Beyond the role of emotion words in emotional development, the connection between language and emotion also manifests in how language development can be influenced by emotionally valenced words. Emotional valence refers to the pleasantness of a word and the extent of its positivity or negativity (Warriner et al., 2013). This broadens the definition of emotion words from just describing an emotional state, such as “happy” or “sad”, to all content words. For example, “holiday” has a positive valence whereas “prison” is negative.

It has been argued that emotional content supports language processing by grounding word meanings in emotional experience. Vinson et al. (2014) explored this idea using lexical decision data from the British Lexicon Project. After controlling for non-emotional variables like word frequency, familiarity, and age of acquisition (AoA), they found that emotionally valenced words are processed faster than neutral ones, regardless of whether the words explicitly refer to emotions. Kuperman et al. (2014) extended these findings using lexical decision data from the English Lexicon Project, confirming that both valence and arousal independently affect word processing. Faster processing of emotional words has been observed across studies using auditory or visual lexical decision (Bayer et al., 2010; Kousta et al., 2009; Ponari et al., 2018). Neuroimaging and electrophysiological studies have shown that words with more extreme valence elicit distinct neural responses compared to neutral words (Pauligk et al., 2019; Scott et al., 2009; Vigliocco et al., 2014; Yao et al., 2016). Even in the context of fluent reading, fixation times on emotional words were faster than those on neutral words (Scott et al., 2012). These findings suggest that the emotional grounding of word meaning facilitates processing, perhaps reflecting the motivational importance of processing emotional stimuli for survival (Kousta et al., 2009).

The facilitatory effect of emotional content may be especially important for the acquisition of abstract words. The affective embodiment account proposes that emotional content helps ground abstract word meanings in emotional experiences, providing a motivational relevance that heightens processing efficiency (Vigliocco et al., 2014). Abstract words that are positively or negatively valenced are rated as being acquired earlier than neutral abstract words (Kousta et al., 2011; Ponari et al., 2018). This

suggests that emotional valence acts as a bootstrapping mechanism for acquiring abstract concepts (Ponari et al., 2018; Vigliocco et al., 2014). Developmental variations exist in this valence effect. Ponari et al. (2018) tested three age groups of 20 children each (aged 6 to 7, 8 to 9, and 10 to 11 years), and they found that the processing advantage of emotionally valenced abstract words was most pronounced in the 8- to 9-year-olds. Lund et al. (2019) tested children aged 5 to 7 years and found a faciliatory effect of valence only in 6- to 7-year-olds. Yet, it is difficult to draw conclusions across these studies due to methodological differences and the different age ranges tested.

Emotional valence also impacts children's learning and memory for newly taught abstract words. Using a word learning experiment, Ponari et al. (2020) found that children aged 7 to 9 years learned emotionally valenced abstract words more effectively than neutral words, and they were more accurate at defining emotionally valenced words too. J. M. Kim et al. (2020) presented 7- to 8-year-old children with spoken word stimuli varying in both valence and concreteness and tested their recognition memory for the words later the same day. They found that children remembered negative abstract words more accurately than neutral and positive ones, but this effect was not observed for concrete meanings. These findings align with the predictions of the affective embodiment account, suggesting that children ground abstract word meanings in their emotional system, with emotional content providing a bootstrapping mechanism for abstract word processing (Kousta et al., 2011). While this account highlights one potential route, namely emotional valence, in language acquisition, other hypotheses emphasise the role of linguistic development and social development (e.g., Borghi et al., 2017; Kousta et al., 2011). The developmental variation reported earlier in Ponari et al. (2018) and Lund et al. (2019) further suggest that different mechanism may vary in importance depending on the child's developmental stages.

Gaps in the literature

Despite considerable research on the close association between language and emotion, several key areas remain underexplored. For one, existing studies often examined emotion words in isolation and primarily used parental reports as measures of emotion understanding. There is limited exploration of

how these words are used in narrative contexts written for and by children. For example, Ridgeway et al. (1985) studied how infants and toddlers (age ranging from 18 to 71 months) used and understood emotion-descriptive adjectives via parental reports. Baron-Cohen et al. (2010) compiled a list of 336 emotion words using an electronic thesaurus and collected parental or teacher reports to chart developmental norms of comprehending these emotion words by children aged 4 to 16 years. While these studies were useful guides to track children's emotion understanding, reports by parents or teachers might be subjective and could not reveal how these words were actually experienced and used by the children themselves in narrative contexts.

Another gap in the existing research on emotion language is that studies have been centred on early childhood, primarily focusing on child-directed speech (Nencheva et al., 2023; Ogren & Sandhofer, 2021). Far less attention has been given to older children, particularly in the context of written language. Spoken language, especially conversational exchanges, is often more immediate and situational, allowing for real-time interactions between individuals. In comparison, written language is more structured and decontextualised, characterised by a richer vocabulary and more complex syntactic structures than everyday conversation (Nation et al., 2022). This language environment is beneficial for vocabulary growth and the development of literacy skills, as it builds the foundation for later language use through reading and writing. The richer language experience provided by books may also afford opportunities for children to encounter more emotion words. Books targeted at preschool children for shared reading contain words that are more emotionally arousing that child-directed speech (Dawson et al., 2021). Given the importance of emotion words in emotional development, it is thus plausible that increased reading helps children acquire more emotion words; and children who read more are more likely to develop a more nuanced understanding of emotions, potentially contributing to better emotional regulation and social interaction (for review, see Batini et al., 2021). However, little research has explored the occurrence and variety of emotion words in children's books.

From the perspective of children's word learning, context provides the environmental and linguistic

backdrop that shapes how words are learned, understood, and integrated into one's vocabulary; and many studies have examined the types of contextual characteristics that facilitate word learning (e.g., Bolger et al., 2008; Chilton & Ehri, 2015; Hulme et al., 2023; Joseph & Nation, 2018; Mak et al., 2021; Pagán & Nation, 2019). However, fewer studies have investigated the affective properties of the context, specifically how the emotional tone of the text may impact word learning through reading (Driver, 2021; Frances et al., 2020; Lana & Kuperman, 2023; Snefjella et al., 2020). Research with adults has shown that emotionally charged narratives can create more memorable contexts for learning, making words encountered in such contexts more salient and easier to retain (e.g., Brierley et al., 2007). To date, only one study has investigated a similar concept in children. Ponari et al. (2020) taught 7- to 10-year-olds new abstract words that were neutral, negative or positive. One group of children were taught the words in a condition that emphasised emotional information, while the other group were taught the new words via more encyclopaedic, non-emotional information. Their results showed that abstract words with emotional valence were better learned, regardless of whether or not the teaching strategy emphasised emotional information. Yet, their study focused on explicit and direct teaching, repeated in multiple sessions extending across a week. This is quite different from how children experience new words in narrative contexts during natural reading. Whether and how valence influences children's learning of new written words via independent reading has not been investigated directly. Understanding how emotional contexts influence word learning as children engage in reading could provide valuable insights for educational practice, especially in creating emotionally supportive learning environments that enhance language acquisition.

Research questions and implications

In light of the existing literature, this thesis aims to examine emotion language in children's reading, writing, and word learning. First, I study the frequency and distribution of emotion words in texts both written by and for children, as well as the emotional valence associated with these words. I then investigate whether the emotional context within narratives influences children's word learning through reading. Finally, I explore how children's experience with reading and written language is

related to the emotional content of the language they produce.

The thesis addresses the following questions:

> Chapter 2 and 3: What emotion words are present in languages written for and by children? How do they compare to child-directed speech and television programmes? How does children's use of emotion words change with developmental age and gender?
>
> Chapters 4 and 5: Can children learn new words via reading emotional narratives? How does emotional narrative context influence novel word learning?
>
> Chapter 6: Does individual differences in reading experience influence variabilities in children's language production, and in particular, production of emotion words?

By identifying the types of emotion words and their emotional valence in children's written language, this research will contribute to a deeper understanding of children's language environment and how emotions are expressed through reading and writing. Moreover, exploring how emotional narratives influence children's ability to learn new words is important because reading is a key mechanism for vocabulary development (Nagy et al., 1985, 1987). Emotional context may create a rich environment for learning by making words more memorable. Furthermore, individual differences in reading experience are likely to contribute to variability in language production, including the use of emotion words. Investigating this relationship will reveal how reading habits influence children's expressive language, especially in the emotional domain.

While the work presented in this thesis is theoretical, there may be broader implications for future translation. Research has shown that children with language or literacy difficulties often face challenges related to mental health or well-being (Boyes et al., 2016, 2018; S. Griffiths et al., 2021; Jelen et al., 2023), though the underlying reason for this association remains unclear. Understanding children's language environment, including how emotions are expressed through reading and writing, can offer insight into what these children might be missing if they are unable to fully engage with

reading. This in turn might help inform the design of literacy-based or alternative interventions to support children struggling with emotional comprehension or mental health difficulties. Exploring how reading habits might influence children's expressive language will provide further evidence to support the national strategy to promote reading and literacy. Additionally, understanding how emotional context supports vocabulary development could have practical implications for educational strategies, as it may guide the design of reading materials and inform teaching strategies to support vocabulary acquisition, particularly for abstract and emotion words.

Scope of the thesis

This thesis explores the relationship between language and emotion, specifically focusing on how emotion words are used and learned in written language. To provide clarity, several key terms are defined in this section.

Emotion language, in this thesis, is defined at a lexical level, referring to individual words that denote emotional content. Previous studies have created lists of emotion words for use in language and emotion research (e.g., Baron-Cohen et al., 2010; Ridgeway et al., 1985; Storm & Storm, 1987). Yet, it is important to note that different sources have employed different definitions of what constitutes an emotion word, and as a result, have included different 'types' of emotion words, from mood-descriptive adjectives to feeling states. Given the lack of consensus in the literature, this thesis approaches emotion words from two levels. At a narrower level, I focus on emotion-descriptive words, which directly label emotional states (e.g., "happy", "sad"), following the set of emotion words identified by Baron-Cohen et al. (2010) and considered in detail in Chapter 2. At a broader level, I focused on emotion-valence words, which encompass all words that convey emotional tone or valence. According to Warriner et al. (2013), valence refers to the pleasantness of a word and the extent of its positivity or negativity. Each chapter will centre on one of these definitions, depending on the focus and methodology.

Emotion language in context refers to the way emotion words are used both in isolation and within the

broader framework of narratives. The thesis specifically focuses on how emotion words are used in written language, as opposed to spoken or conversational language. Relatedly, the thesis also examines how the emotion provided by the context influences the learning of new words via reading.

Finally, it is important to clarify what the thesis will not cover. For one, this thesis does not explore the distinction between knowing an emotion word and fully understanding or expressing the associated emotion. While children may learn the word “happy”, this does not guarantee that they fully comprehend the emotional experience of happiness or can express it in a socially appropriate way. The focus remains on the use of emotion words in language, not on whether children or adults can accurately recognise or understand the emotional states these words describe, or how the use of emotion words contributes to emotional recognition and regulation abilities.

Second, although the dimensional approach to emotion includes both valence and arousal, this thesis focuses on valence—the emotional positivity or negativity of a word. Emotional arousal, which refers to the intensity of emotion, is not a primary focus of this research. While Chapters 3 and 5 report computational measures related to arousal, these measures are not interpreted in the context of the main research questions.

Thesis structure

My DPhil thesis examines the overarching theme of emotion language from multiple angles, through a combination of data-driven approaches and experimental designs. The thesis is organised into seven chapters, each focused on a different aspect of emotion language. This chapter, Chapter 1 introduces the broad topics of language and emotion, setting the stage for the specific investigations that follow. The five content chapters 2 to 6 have their own introduction, methods, results, and discussion for the specific topics explored in each chapter. Chapter 7 reviews and discusses the findings of all chapters and reflects on common themes throughout the thesis.

Chapters 2 and 3 employ data-driven corpus analyses, where I examine emotional language in various corpora of children’s language. Chapter 2 focuses on a specific set of emotion words (Baron-Cohen et

al., 2010), charting the frequency and diversity of these words across children's language environments. Chapter 3 zooms in on children's own writing, applying sentiment analysis to explore the emotional content of free narrative writing and its development over time and differences between genders. By analysing naturally occurring language in children's written materials, this method offers insights into the patterns and contextual usage of emotion words. This approach provides a deeper understanding of how emotional language is integrated into children's writing and language environments.

Chapters 4 and 5 shift focus to an experimental approach, reporting on word learning experiments that directly assess the influence of emotional narrative context on the learning of novel adjectives. These experiments explore how exposure to emotionally charged narratives impacts memory for words both in adults (Chapter 4) and children (Chapter 5). The experiments provide insights into the causal relationships between emotional context and word learning, offering a more controlled perspective on how emotion language develops. Chapter 5 also incorporates a computational analysis that quantifies the emotional context of words within a children's writing corpus using valence norms (Warriner et al., 2013). Finally, Chapter 6 integrates multiple approaches of the thesis to extend the corpus findings and to investigate whether individual differences in children's reading experience, measured as print exposure, is associated with their production of emotion language.

Chapter 2. Emotion Words in Children's Written Language

This chapter is based on the submitted preprint:

Dong, Y., & Nation, K. (2024). *Charting the frequency and diversity of emotion words in children's language: Written language matters*. Submitted to First Language (19th September 2024); Major revision decision (29th October 2924); pre-print available at: OSF. https://doi.org/10.31219/osf.io/zj5k7

Chapter summary:

Emotion words allow us to identify, describe, and regulate our emotional states. Emotion vocabulary grows through childhood, but little research has considered emotion words in the context of children's written language. To address this gap, we used a cross-corpus developmental approach to chart the emergence of emotion words in children's reading experience, and in their own writing. For comparison, we also captured occurrences of the same set of emotion words in age-matched samples of children's spoken language experience via caregiver child-directed speech and television programmes. We observed that even books targeted at pre-schoolers for shared reading contained more unique emotion words than both caregiver speech and television language. As the targeted age of books increased through mid-childhood and early adolescence, the frequency and diversity of emotion words increased further. This pattern was also seen in children's own writing, with more unique and diverse emotion words being used by older children. These findings indicate that written language requires children to comprehend and produce emotion words that are rare in everyday conversations. We speculate that this linguistic experience may play a role in emotional development by providing opportunities to consider and communicate mental situations beyond the everyday.

Introduction

Language is considered a "key ingredient" in the development of social-emotional understanding and empathy (Lindquist, 2017). Language offers the tools for identifying and describing our own

emotions and those of others. From the early stages of development, children's linguistic experiences are deeply embedded in social interaction, setting the stage for their burgeoning emotional understanding later on (Dunn et al., 1991). There is good evidence to show that emotion words, or affective labels, play a fundamental role in emotional development. For example, the breadth of pre-schoolers' emotion vocabulary has been linked to their emotion regulation strategies (Streubel et al., 2020), and more generally, the capacity to label and differentiate emotions has implications for mental well-being through childhood, adolescence, and adulthood (e.g., Nook, 2021; Torre & Lieberman, 2018). Most studies on the development of emotion vocabulary have focused on infancy and early childhood, and children's experience with emotional language has been captured via child-directed speech, typically sampling the language of caregivers as they interact with their children (e.g., Nencheva et al., 2023; Ogren & Sandhofer, 2021). There has been less emphasis on mid-childhood, yet Baron-Cohen et al. (2010) found that the size of children's emotion word vocabulary doubles every two years between 4- and 11-years of age. Interestingly, this expansion in emotion word vocabulary aligns with the onset of literacy. Children's books contain more words and more diverse and complex words than child-directed speech (e.g., Dawson et al., 2021) and as children learn new vocabulary via reading (e.g., Joseph et al., 2014), it is likely that emotion words, especially more complex emotion words, might be acquired via experience with written language. This feels a plausible hypothesis, but lacking from the evidence base is any information about the nature of emotion vocabulary afforded by written language input. This is a critical gap in our understanding of how emotional development might be influenced by written language. We address this gap in two ways. First, we chart the occurrence of emotion words in books targeted at children of different ages and second, we consider children's usage of emotion words in their own writing in cross-sectional analyses of stories written by 5–13-year-olds. Our aim is therefore to describe the occurrence and usage of emotion words in language written *for* children and in the language written *by* children themselves. We also ask how this compares with occurrence of the same set of emotion words in child-directed speech, and in the language of television shows targeted at children of the same age.

According to the constructionist hypothesis of emotional development, language supports the

acquisition of emotional knowledge and is considered critical to the development of different aspects of emotional understanding (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017; Shablack & Lindquist, 2019). Language is situated in social interaction from infancy onwards and this is reflected in the association between children's experience with emotional language and their subsequent emotional understanding. For example, Dunn et al. (1991) coded "feeling-state" language (e.g., happy) from recordings of everyday mother-child interactions at 36 months of age. They noted that children who experienced more feeling-state language at 36 months were better able to understand others' emotions at 6 years of age. In line with the view that language is critical to the construction of emotion, there is also an association between language competence and emotion understanding in pre-schoolers (e.g., Widen & Russell, 2003, 2008) and in older children (e.g., Beck et al., 2012; S. Griffiths et al., 2020), with broader implications for mental health (e.g., Jelen et al., 2023; Wilmot et al., 2023).

Research examining how young children experience and use emotion words within their language environment has provided valuable insights into the linguistic underpinnings of emotional development. Contributing to this, Ogren and Sandhofer (2021) analysed language transcripts from the Child Language Data Exchange System (CHILDES) database for children aged 15 to 47 months. They described the characteristics of young children's emotion lexicon, including the frequency and variety of emotion words and factors influencing individual differences in their production. Nencheva et al. (2023) used data from Wordbank (Frank et al., 2017; infant receptive and expressive vocabulary as estimated by the MacArthur-Bates Communicative Development Inventory parental checklist) and CHILDES (caregiver speech) to study the relationships between toddlers' language production and caregiver input. They found that toddlers develop a broad emotional vocabulary early in life, influenced by a network of semantically connected words. They also showed in a longitudinal analysis that how caregivers use emotion and mental state labels predicts children's emotion label production overtime.

While these studies document important links between language input and the development of

emotional vocabulary in children, the field has primarily focused on the spoken language environment, especially within the context of child-directed speech and early childhood interactions. Yet, emotional vocabulary continues to develop beyond early childhood and might therefore be facilitated by the onset of literacy. Baron-Cohen et al. (2010) used parent/teacher checklist data to estimate whether 4-16 year-olds understand 336 emotion words. They found a general increase in emotion word vocabulary by age, and they noted that emotion word vocabulary doubles every two years between 4- and 11-years of age. In line with this observation, Ponari et al. (2018) analysed age of acquisition ratings for over 13,000 words and identified 8.5 years as the age which shows the steepest increase in acquisition of abstract words, including words describing emotions. These findings are certainly consistent with the idea that written language provides critical language input that might help drive the acquisition of emotion vocabulary. Reviewing the nature of children's "book language", Nation et al. (2022) argued that exposure to written language provides children with opportunities to experience language that is different to everyday usage, including the complex and nuanced language associated with emotions and mental states. Even books written for pre-schoolers contain more linguistically complex words than child-directed speech (Dawson et al., 2021; Montag et al., 2015). For older children, Korochkina et al. (2024) found that 28% of words in books for 7- to 9-year-olds never appear in age-appropriate television programmes. In addition, Dawson et al. (2021) noted that books for pre-schoolers contain more emotionally arousing words than child-directed speech. Readers need to navigate book language to understand social relationships and make inferences about the mental state of characters (Y.-S. G. Kim et al., 2015; Mar et al., 2006; Siller et al., 2014; White et al., 2009) and children's story books therefore provide opportunities to adopt the emotional perspectives of different characters and to understand empathy (e.g., Hogan, 2011; Kucirkova, 2019). Being read to, joint reading, and independent reading are all associated with children's social, emotional, and educational outcomes (Green et al., 2023). From this view, the development of emotion word vocabulary in mid-childhood might be a result of experience with books and reading. Thus, as children's language experiences evolve with the onset of literacy, there is a clear need to investigate how written language affords opportunities for children to experience emotion words.

One way to measure children's understanding and use of emotion words is to elicit them in production tasks. Several studies have reported that children's production increases through childhood and becomes more nuanced and multidimensional over time (e.g., Grosse et al., 2021; Nook et al., 2020). For example, emotion word labelling of facial expressions increases with age between 3 and 5 years (Widen & Russell, 2008). Using a vignette test which depicts a child protagonist in a typical emotion-eliciting situation, Grosse et al. (2021) asked 123 4- to 11-year-old children to name the emotion the protagonist in the vignette might feel. They found that older children produced more emotion words, and that their usage became more adult-like. Using a word generation task (e.g., "think about a child who is feeling sad and then write down as many words as you can to describe this feeling"), Doost et al. (1999) found that older children in secondary schools produced significantly more emotion words than the younger children in primary schools. While these findings mirror the increases in emotion word comprehension through mid-childhood reported by Baron-Cohen et al. (2010), it is notable that existing studies have all probed emotion word usage directly, using simple word generation or picture prompts. Lacking is a more nuanced and naturalistic perspective on children's use of emotion words.

In summary, the importance of emotion words in emotional development is well accepted. Lacking, however, are data that consider emotion vocabulary in children's written language experience, both in terms of what children read, and what they produce themselves. We took a corpus linguistics approach to address these two evidence gaps, using data from several corpora that sampled language targeted at children of different ages via child-directed speech, television shows and books, and one corpus of stories written by children. Across each corpus, we considered the 336 emotion words identified by Baron-Cohen et al. (2010) and described the frequency and diversity of these emotion words. We predicted that children's books would contain a more extensive and diverse range of emotion words than child-directed speech and television, and we also expected language written for older children to contain more diverse use of emotion words than younger children, and for this pattern to also be seen in children's own writing.

Method

We used six children's language corpora (Table 2-1) and investigated the relative occurrence of 336 emotion words, as selected by Baron-Cohen et al. (2010). They defined an emotion word as one that describes a mental state with an emotional dimension (i.e., it could be preceded by 'I feel x' or 'he/she looks x'). The full set of 336 words are reproduced in Appendix 2, along with item-level summary statistics from our analyses.

Table 2-1. Description of the six corpora used in this paper.

Corpus name	Source	Targeted Age Range	Size (Type)	Size (Tokens)
Child-Directed Speech	CHILDES-UK (MacWhinney, 2000)	0-6 years	24,129	3,853,976
Picture Books	ReadOxford (Dawson et al., 2021)	0-7 years	11,561	319,435
CBeebies (TV)	SUBTLEX-UK (van Heuven et al., 2014)	0-6 years	27,236	5,848,083
CBBC (TV)	SUBTLEX-UK (van Heuven et al., 2014)	6-12 years	58,691	13,612,278
Reading Books	Oxford Children's Language Corpus -- 2019 Reading (OUP)	5-16 years	377,108	63,494,697
Children's Writing	Oxford Children's Language Corpus -- 2019 Writing (OUP)	5-13 years	223,594	47,841,388

Note: OUP stands for Oxford University Press. See Appendix 2 for the age distribution for all tokens in child-directed speech and picture books.

Child-directed speech was generated from 10 corpora in the English-UK section of the CHILDES database (MacWhinney, 2000). The sample comprised all suitable corpora from this collection, except

for those that focused on specific populations (e.g., children with language impairments). The final set of 10 corpora contained transcripts of interactions between 190 different children aged 6 weeks to 6 years and their caregivers, siblings, other family members and research investigators. Recordings took place across a variety of contexts, but typically involved structured and free play activities between children and their caregivers, as well as everyday routines such as mealtimes and bedtimes. Across all recordings, utterances produced by the child were filtered out, such that the final dataset contained only talk directed to the child for a total word count of 3,853,976. See Appendix 2 for the age distribution for all tokens in the child-directed speech.

The *Picture Book corpus* comprised 160 children's fiction books, selected to be representative of the type of reading material children encounter in shared reading contexts in the UK (Dawson et al., 2021). This captures book language input that is age-matched to the child-directed speech documented in CHILDES. The corpus was created by first generating a list of titles with a target age range of 0-7 years from a combination of retailer bestseller lists and recommendations from literacy charities, book review sites, and teachers. The final list included the titles that were cited most frequently across a combination of retailer bestseller lists and recommendations from literacy charities, book review sites, and teachers. The dataset contained 319,435 words. See Appendix 2 for the age distribution for all tokens in the picture books.

The *Reading Book corpus* was sampled from the reading component of the Oxford Children's Language Corpus developed and held by Oxford University Press (OUP, Wild et al., 2013). It was initiated in 2006 to guide the preparation of children's dictionaries. The corpus contains a wide range of material that children encounter during reading experience, including classic and modern fiction, non-fiction, textbooks, websites, and magazines. We accessed the 2019 version of the corpus, totalling 63,494,697 words written for 5- to 16-year-old children. Frequency lists of all the words in the corpus was downloaded from Sketch Engine (http://www.sketchengine.eu), with the kind permission of Oxford University Press. Some documents (45.8%) were tagged with targeted Key Stage. This refers to bandings within the education system of England and Wales, with 5-7-year-olds falling into Key

Stage 1, 7-11-year-olds into Key Stage 2, 11-14-year-olds into Key Stage 3, and 14-16-year-olds into Key Stage 4. Books targeting children at Key Stage 2 were most prevalent (18 million tokens), followed by Key Stage 3 (14 million tokens). Books targeting younger children (Key Stage 1) and older children (Key Stage 4) are smaller, each comprising 2 million tokens.

CBeebies and *CBBC* data were extracted from the children's sub-corpora of SUBTLEX-UK (van Heuven et al., 2014), which is a corpus of subtitles from television programmes shown by the British Broadcasting Corporation (BBC). CBBC content is targeted at children aged 6-12 years, while its sister channel, CBeebies, is aimed at pre-school children. The CBeebies norms were derived from 5,848,083 words and contain 27,236 types, while the CBBC norms were derived from 13,612,278 words and contain 58,691 types, as reported in van Heuven et al. (2014). We used these data as a proxy for spoken language input, but note that the discourse and narrative context of television language is quite different to the type of spoken language experienced in day-to-day conversation, as indexed by CHILDES and child-directed speech. We return to consider this further in the discussion.

The *Writing Corpus* was sampled from the writing component of the Oxford Children's Language Corpus developed and held by Oxford University Press (Wild et al., 2013). This is a dynamic corpus that contains stories submitted as part of the BBC 500 Words annual writing competition for children. This has been running for over 10 years with the same format in which children (aged 5-13 years until 2020, and then aged 5-11 when the competition returned in 2023) across the UK were invited to submit a story on any theme or topic, so long as the word count was not greater than 500 words. We selected all stories submitted in 2019 (N = 107,273; approximately 55-million word tokens; see also Dong, Hsiao, et al., 2024; Hsiao et al., 2023, 2024). Each story was tagged with the Key Stage of the child author. Most entries (59%) came from children in Key Stage 2; 39% of entries came from children in Key Stage 3 and only 2% from the youngest children in Key Stage 1. More girls (59.44%) contributed stories than boys (40.56%).

Results

We analysed the distribution and diversity of the 336 emotion words identified by Baron-Cohen et al. (2010) across each corpus and our results are organised into three short sections. First, we considered the language directed at pre-school children and compare emotion word input in picture books, child-directed speech and television subtitles from CBeebies. We then considered language targeted at older children and compared emotion vocabulary in children's reading books with television subtitles from CBBC. Finally, we considered children's own production by analysing the occurrence of emotion words in their own written stories. Data and code associated with this paper are available on Open Science Framework: http://tiny.cc/rainy_thesis_chapter2.

Given the size of the datasets and the potential issue of overpowering (Egbert et al., 2022), inferential statistics are not necessarily helpful. We thus describe and interpret patterns based on cross- and within-corpus visualisations. To visualise the distribution of all emotion words in each corpus, we calculated the number of occurrences of each emotion word on the list in the whole corpus and plotted a frequency graph sorting the emotion words by their frequency. To compare the diversity of emotion words between corpora of different sizes, we adopted the sampling methods used by Montag et al. (2015) and Dawson et al. (2021). We took multiple random samples from each corpus, ranging from 100 to 50,000 words in size, and increasing in increments of 100 words each time. One hundred simulations were generated at each sample size, each based on a new random sample. We then aggregated and plotted the number of emotion word types at each sample size.

(i) Language directed at pre-school children

Figure 2-1 (panels a-c) shows the overall distribution of the emotion words in the three corpora (child-directed speech, CBeebies, picture books). The x-axis is frequency rank of individual words from most frequent (left) to least frequent (right). The y-axis is the raw frequency of the emotion word. We see that all distributions are Zipfian, where the frequency of a word is inversely proportional to its frequency rank. We observed that spoken language in child-directed speech and children's television programmes has a steeper slope and a longer tail, while written language in children's picture books is more spread.

Figure 2-1. Frequency of emotion words in language for younger children (a-c) and older children (d-e).

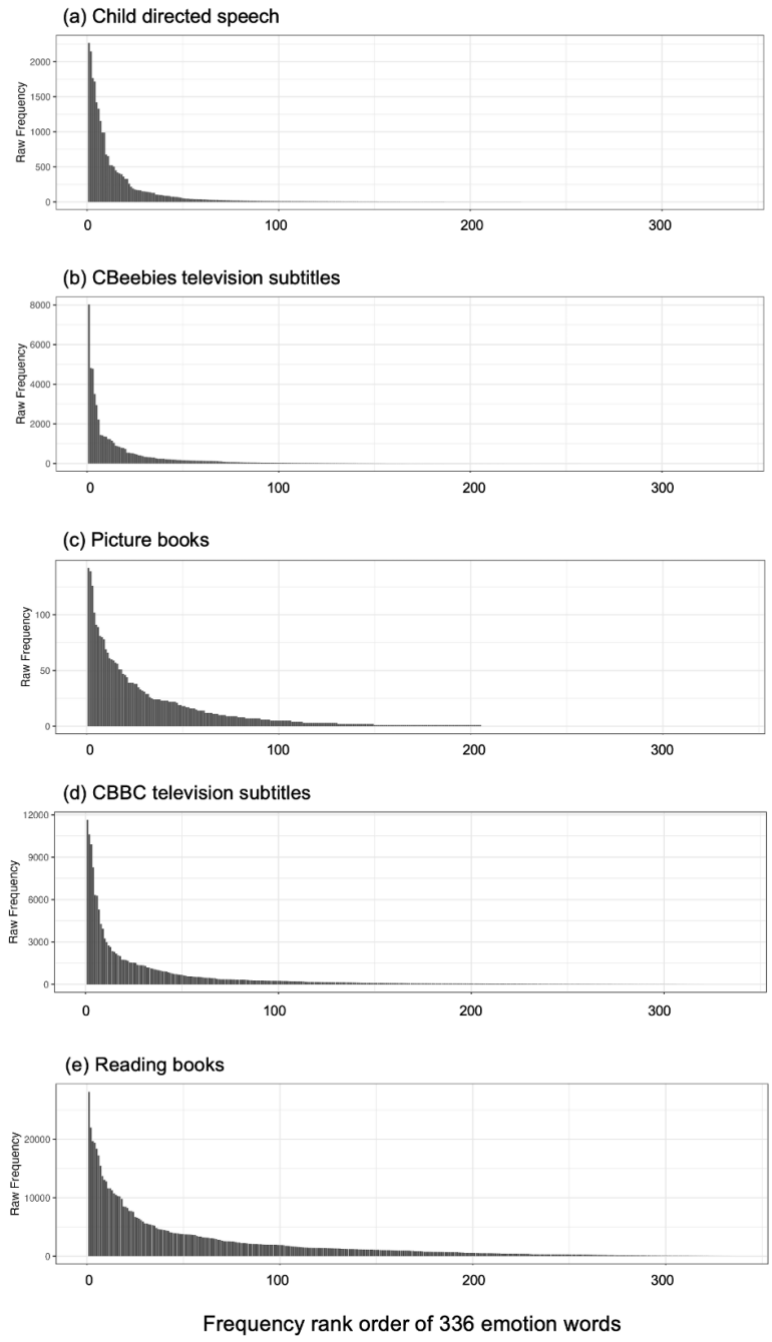
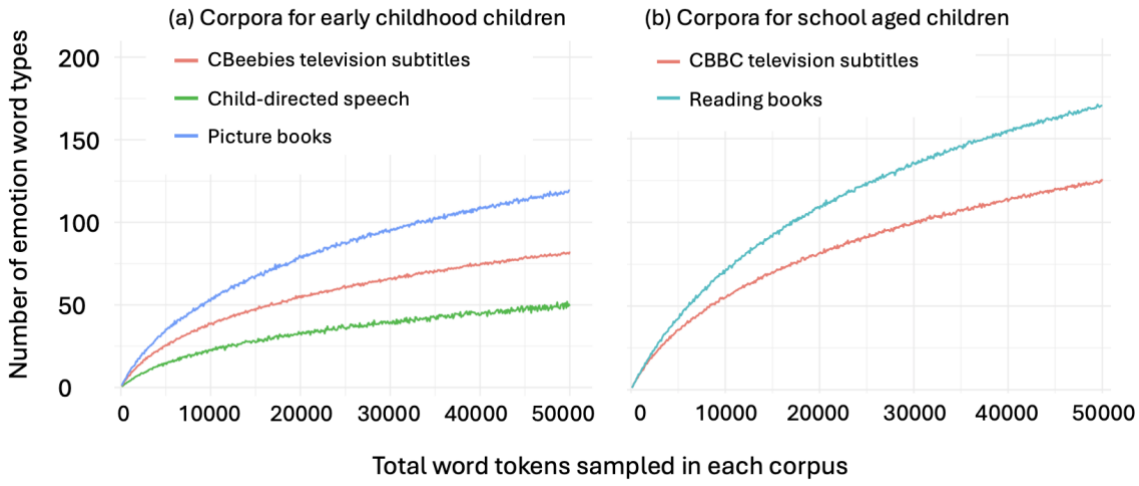


Figure 2-2 (panel a) shows the mean number of emotion word types at each sample size in the same

three corpora, namely child-directed speech, CBeebies and picture books. The data show that at each sample size, the picture book corpus contains a greater number of unique emotion word types than CBeebies subtitles, followed by child-directed speech. Differences also emerge in the slopes of the lines. The number of emotion word types in picture books shows a greater increase in unique word types per unit increase in word token than both CBeebies and child-directed speech.

Figure 2-2. Diversity of emotion words in language for (a) younger children and (b) older children.

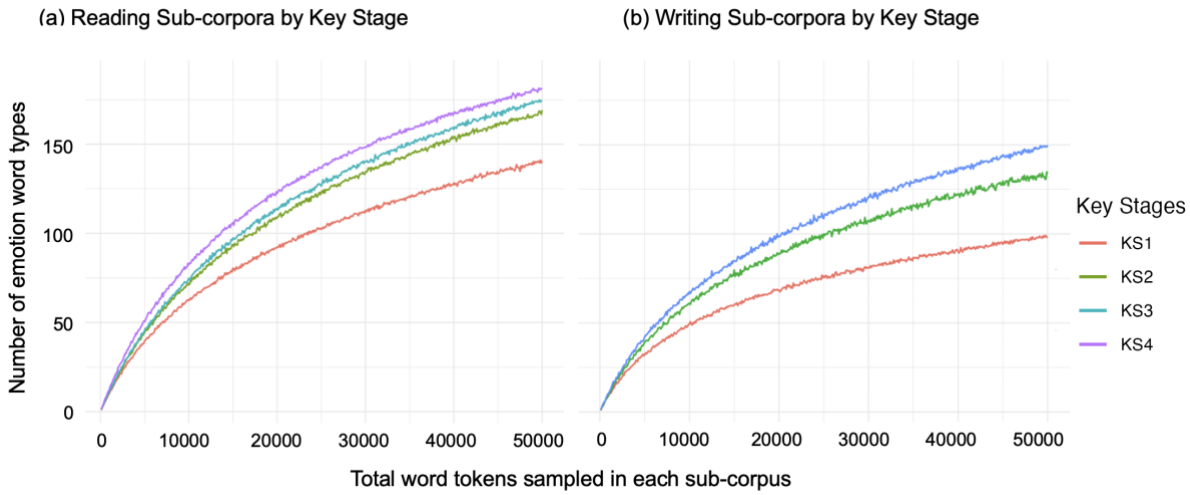


(ii) Language directed at older children

Figure 2-1 (panels d-e) shows the frequency distributions of emotion words in the CBBC and reading corpora, which once again are clearly Zipfian. Emotion words in CBBC television subtitles show a steeper gradient and a longer tail, while written language is more spread, with less concentration of high frequency words. Turning to the lexical diversity of the two registers, Figure 2-2 (panel b) shows that at any given sample size, books contain a greater number of unique emotion word types than CBBC television subtitles. The steeper gradient for book language indicates a greater increase in unique word types per unit increase in word tokens.

To model lexical diversity across targeted age, we used the Key Stage metadata that was available for around half of the reading books. We conducted sampling at each Key Stage to calculate the unique number of emotion word types at random samples of each corpus. Figure 2-3 (panel a) shows that at any given sample size, books targeted at older children contain a greater number of unique emotion word types than books targeted at younger children.

Figure 2-3. Diversity of emotion words in children's reading and writing by Key Stage.



(iii) Language written by children

Each story in the children's writing corpus was tagged by the Key Stage of the child author. Most entries (59%) came from children in Key Stage 2; 39% of entries came from children in Key Stage 3 and only 2% from the youngest children in Key Stage 1. We conducted sampling at each Key Stage to calculate the unique number of emotion word types at random samples of each reading corpus. Figure 2-3 (panel b) shows that at any given sample size, stories written by older children contain a greater diversity of emotion words than stories written by younger children.

Discussion

Our study charted emotion word frequency and diversity across children's language, with a focus on written language. Comparing language input afforded by the language children hear with the language they read (and hear in the context of shared reading), our findings show a greater diversity of emotion words in book language than spoken language. The frequency distribution of emotion words was Zipfian, mirroring the general distribution of word frequencies in language, where a small set of words appear very frequently, with the majority forming the long tail of the distribution (Piantadosi, 2014). As compared to spoken language, the frequency distribution for emotion words in books is more spread, suggesting that children are more likely to encounter low frequency emotion words in books rather than via child-directed speech or from watching television. In language targeted at pre-school children, picture books contained a greater diversity of emotion words than the language of CBeebies television programmes, which in turn, contained more diversity than child-directed speech. Similar patterns were observed for older children, where books for independent reading contained more emotion words than the CBBC television programmes. We were also able to compare books by Key Stage, and this showed that books written for older children contain a wider range of emotion words relative to those written for younger children. Mirroring this, older children used more emotion words in their own writing, relative to the stories written by younger children.

At a general level, our findings are consistent with other investigations of children's book language that document greater linguistic complexity and diversity in books than child-directed speech (Dawson et al., 2021; Hsiao et al., 2023). These differences can be attributed to the inherent characteristics of written versus spoken language. Written language, being more decontextualised, requires readers to rely on the text itself to understand the intended meaning, without the aid of paralinguistic cues and a shared situation. This necessitates lexical diversity, and greater linguistic complexity more generally. Clearly, spoken language interactions provide rich input from which children can learn about emotion via shared context and multimodal cues such as prosody, facial expression and gestures. When reading, however, the greater diversity of emotion words in written

language exposes children to a wider range of nuanced emotion words and words that are rarely encountered in everyday conversations.

Extending beyond previous work that has compared lexical diversity in children's books vs. child-directed speech, our study included comparisons with television language. We found that both CBeebies and CBBC programmes contained a greater range of emotion words than child-directed speech, but both had less diversity than books. This middling position fits with the observation that while television programmes are often scripted, they also contain elements of spontaneous speech (Y. Zhang & Gu, 2023). Similar findings have been reported for other types of media language, such as the language used in educational apps for pre-schoolers. Kolak et al. (2023) found these contained lower frequency words than child-directed speech yet shorter utterances than books. Gowenlock et al. (2024) conducted a scoping review looking at the exposure to language in video and its impact on linguistic development in children aged 3 to 11 years, and they found that the results are mixed, which depends on the video quality and the viewing context. Looking across the different input corpora in our study, our findings suggest that engaging with books allows children to encounter more diverse language and language experience that is distinct from everyday speech and other types of media, including emotional language. Books therefore offer children the chance to encounter language that portrays emotion and mental states and in turn, this might support the development of a more comprehensive understanding of emotional concepts.

Written language aimed at older children contained a more diverse range of emotion words than books written for younger children. In particular, emotion word diversity began to increase for 7–11-year-olds, as estimated by Key Stage 2 books. This was reflected in children's own writing, where older children produced more emotion words, and the difference was largest between Key Stage 1 and Key Stage 2. This echoes with Baron-Cohen et al.'s (2010) parent checklist data, which estimated a rapid expansion in emotion word vocabulary at about 8.5 years. We suggest that as children enter Key Stage 2 and reading expertise expands, books offer an opportunity to experience rich and nuanced language, including words that articulate emotions and mental states. This experience might then

serve to enrich children's understanding of emotion concepts, such that the increased intake of emotional language is also reflected in how children use emotion words in their own writing. This is supported by the two similar looking graphs in Figure 2-3. Indeed, the act of crafting writing might itself be a driver for developmental change. On this view, the need to produce written language in such a way that it conveys the intended meaning for the reader might extend children's use of emotion language to communicate empathy and emotional states, and in turn, this will be reflected in greater understanding of empathy and emotional states (see Dong, Hsiao, et al., 2024, for further discussion). We are not able to forge direct links across the two corpora and so this speculation awaits testing experimentally, and longitudinal data would be particularly valuable. It is also important to note that using an emotion word does not guarantee that a child understands or can express the associated emotion, and vice versa. Previous research by Ogren and Sandhofer (2021) and Nencheva et al. (2023) studied how children's spoken production of emotion and mental state labels is scaffolded by caregivers' use of these labels in early childhood. Future research could extend these approaches to older children, and to written language, and assess how a child's written and spoken language experience is related to their comprehension and production of emotion words in both speech and in writing. Findings from such a study could provide valuable and more direct insights into how language experience shapes emotional development, and how language experience provides the linguistic proficiency needed to express nuances in emotion.

Our intention was to document the occurrence of emotion words by corpus and from this make inferences about when children experience (and produce) emotion vocabulary. It is important to note that while this 'bag-of-words' approach captures type and token frequency and reveals patterns that can inform psychological theory (e.g., Jackson et al., 2022, p. 807), it does not capture the language environment in which words occur. This is important as contextual factors such as negation and figurative usage (e.g., metaphors and idiomatic expressions) can alter the meaning of emotion words. Linguistic context in written language is important, not least because of the decontextualised nature of reading. In child-directed speech and shared reading, for instance, adults might describe emotions about specific incidents, events, or narratives. This interaction does more than introduce children to

emotion words; it contextualises them, and potentially enriches children's understanding of emotional concepts. The language used by adults to describe these situations often mirrors the valence of the emotion words themselves, providing a comprehensive emotional landscape (e.g., Nencheva et al., 2023). For example, words accompanying a description of a joyful event are likely to be positively valenced, thereby reinforcing the emotion concept of happiness. Moreover, the diversity of word types in children's language input is not just related to the amount of language that children are exposed to in different modalities, but is also associated with variables such as the size of caregivers' vocabularies (Green et al., 2023; Montag et al., 2015). While children surely build from this foundation as they read, the act of independent reading will require them to access new language from the books directly, and to do so without the scaffolding support of a caregiver or teacher. The richness and complexity of book language provides opportunities to learn this language, and future work should harness the power of large language models operating over different types of language corpora to take a data-driven approach to understanding emotion language and its emergence in language written for and by children, beyond investigation of a set of words pre-selected as being "emotion words".

Another limitation we must acknowledge is that our results are based on the analysed corpus, and interpretations should consider the specific characteristics of this dataset. Child-directed speech from CHILDES primarily reflects language children hear in the home environment, whereas young children also receive significant speech input from their peers and preschool settings, which are not captured by the current dataset. Moreover, while both the child-directed speech corpus and the picture book corpus were aimed primarily at preschool-aged children, the age distributions of tokens within these two corpora differs (see Appendix 2). Adopting the lower bound of the target age range for the picture books, the token distribution skews toward older children, with the majority of tokens aimed at children aged four and above. This is understandable as picture books targeting older children are generally longer and contain more words. Together, these differences between with the corpora could introduce confounds when interpreting the greater diversity of emotion words in the picture book corpus compared to child-directed speech, which should be further tested in future research. Despite

these limitations, our study represents a valuable first step in identifying trends in children's language exposure. Future research could provide a more accurate depiction by taking recordings from a variety of settings such as in schools and with other stakeholders to better understand children's language environment.

In conclusion, while previous work has considered emotion word development, the potential role of experience with book language has not been considered directly. This is unfortunate as written language has discourse properties that might provide ideal opportunities for children to experience and therefore learn emotion words, and perhaps nuances in emotion concept too. To start to fill this gap, we charted the frequency and diversity of emotion words across different language corpora documenting children's books, child-directed speech and television media, as well as the emotion words children use in their own written stories. We found a consistent pattern of more diverse emotion words in children's books. This is consistent with children encountering more complex emotion words and mental state labels via reading, which subsequently becomes reflected in their own writing.

The next chapter zooms into children's own writing to reveal patterns of language use beyond the individual words, considering the overall direction of the language via emotional valence.

Chapter 3. The Emotional Content of Children's Writing

This chapter is based on the published paper:

Dong, Y., Hsiao, Y., Dawson, N., Banerji, N., & Nation, K. (2024). The Emotional Content of Children's Writing: A Data-Driven Approach. *Cognitive Science*, *48*(3), e13423. https://doi.org/10.1111/cogs.13423

Chapter summary:

Emotion is closely associated with language, but we know very little about how children express emotion in their own writing. We used a large-scale, cross-sectional, and data-driven approach to investigate emotional expression via writing in children of different ages, and whether it varies for boys and girls. We first used a lexicon-based bag-of-words approach (after Hipson & Mohammad, 2020) to identify emotional content in a large corpus of stories (N>100,000) written by 7- to 13-year-old children. Generalized Additive Models were then used to model changes in sentiment across age and gender. Two other machine learning approaches (*BERT* and *TextBlob*) validated and extended these analyses, converging on the finding that positive sentiments in children's writing decrease with age. These findings echo reports from previous studies showing a decrease in mood and an increased use of negative emotion words with age. We also found that stories by girls contained more positive sentiments than stories by boys. Our study shows the utility of using large-scale data-driven approaches to reveal the content and nature of children's writing. Future experimental work should build on these observations to understand the likely complex relationships between written language and emotion, and how these change over development.

Introduction

As a form of expressive language, writing is a complex skill that draws on an array of cognitive factors from choosing words, constructing sentences and building paragraphs, through to developing ideas that are appropriate for the discourse context and genre (Graham, 2006; Harris et al., 2009). For

children, learning to write is difficult and takes time (e.g., Graham, 2018; McCutchen, 2006). In this paper, we focus on one important component of children's writing, namely their use of language to express sentiments and emotional content. Our study builds from the hypothesis that emotional understanding is constructed through development, and that language plays a key role in this construction (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017). While this body of research has considered language in general, our focus is with written language – the language that children read or write. Analyses of the language content of children's books demonstrate that they are more linguistically complex than everyday conversational language, and that their emotional content is more nuanced (Dawson et al., 2021; Dong & Nation, 2024; Nation et al., 2022). If language is a key ingredient in the development of different aspects of emotional understanding (for review, see Shablack & Lindquist, 2019), it follows that experience with *written* language might be particularly associated with the construction of emotion through development. With this as a backdrop, we used natural language processing to analyse a large corpus of stories written by 7- to 13-year-old children. This data-driven cross-sectional approach allowed us to identify emotional content in the narrative writing of girls and boys at different ages.

Language and emotional development

According to the constructionist hypothesis of emotional development, language supports the acquisition of emotional knowledge and is considered a "key ingredient" in the development of different aspects of emotional understanding (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017; Shablack & Lindquist, 2019). Language is situated in social interaction from infancy onwards and this is reflected in the association between children's experience with emotion language and their subsequent emotional understanding. For example, Dunn et al. (1991) coded "feeling-state" language (e.g., happy) from recordings of everyday mother-child interactions at 36 months of age. They noted that children who experienced more feeling-state language at 36 months were better able to understand others' emotions at 6 years of age. In line with the view that language is critical to the construction of emotion, there is also an association between language competence and emotion

understanding in preschoolers (e.g., Widen & Russell, 2003, 2008) and older children (Beck et al., 2012; S. Griffiths et al., 2020).

Emotion vocabulary grows through childhood and becomes more nuanced and multidimensional over time (e.g., Grosse et al., 2021; Nook et al., 2020). Using a word generation task (e.g., "think about a child who is feeling sad and then write down as many words as you can to describe this feeling"), Doost et al. (1999) found that older children produced significantly more emotion words than the younger children. This developmental period coincides with the onset of literacy and this might be important. Nation et al. (2022) argued that exposure to written language provides children with opportunities to experience language that is different to everyday usage, including the complex and nuanced language associated with emotions and mental states. They suggested that the language needed to drive emotional development through childhood might build directly from experience with written language. Consistent with this, Dawson et al. (2021) found that story books written for preschoolers (3-5 years old) contained more emotionally arousing words than child-directed speech. From mid-childhood onwards, most new vocabulary is learned via reading suggesting that the doubling of emotion word vocabulary seen every two years between 4- and 11-years of age (Baron-Cohen et al., 2010) is associated with reading experience. From this perspective, the extended development of connections between emotion and language (e.g., Grosse et al., 2021; Nook et al., 2020) reflects the time needed for these connections to accumulate from written language input. Reflecting on the importance of words in shaping the perception of emotion in faces, Doyle and Lindquist (2018, p.62) noted that the "scowls, frowns, grimaces and growls you see over time presumably develop into conceptual knowledge for what *anger* looks like, helping you to make meaning of new instances of facial actions as instances of anger". In the absence of an actual face, written language needs to describe the face so that the reader can infer its emotional state, as intended by the writer. Consequently, *scowl*, *frown*, *grimace* and *growl* are specified in written language but might not be needed in a conversational exchange where the shared situation can provide the communicative intent, perhaps via facial expression or tone of voice. This partly explains why written language tends to be more linguistically complex than speech (Dawson et al., 2021; Nation et al.,

2022), and indicates why written language (via reading or writing) might provide opportunities for children to experience emotional and mental state language (Dong & Nation, 2024), and in doing, drive their emotional development. Readers are required to navigate this complex language to understand social relationships and make inferences about the mental state of characters (Y.-S. G. Kim et al., 2015; Mar et al., 2006; Siller et al., 2014; White et al., 2009). Children's story books therefore provide opportunities to adopt the emotional perspectives of different characters and to understand empathy (e.g., Hogan, 2011; Kucirkova, 2019). Over time, this might build to the complex nexus between print exposure, literacy, fiction reading, mentalising, emotion understanding, and theory of mind seen in older children and adults (e.g., Beck et al., 2012; Boerma et al., 2017; Kidd & Castano, 2013; Mar & Oatley, 2008; Schwering et al., 2021).

Writing and emotional development

As children learn to read and experience text as a reader, they also begin to write. The act of writing provides opportunities for children to use emotion language in ways that are different to their everyday spoken interaction. In the absence of situational cues, children need to consider the emotional perspectives and mental states of characters and situations they wish to portray, and to nuance their language to communicate this intended meaning to the reader. Potentially then, learning to write and writing experience might drive aspects of emotional development, and vice versa. While the relationship between *reading* and emotional development has been studied in the context of theory of mind and the Event Indexing Model (e.g., Kidd & Castano, 2013; Mar & Oatley, 2008; Zwaan et al., 1995), the possible relationship between children's own *writing* and their emotional development has not been considered.

There are, however, reasons to consider production via writing as important, beyond reading experience itself, and this suggestion is consistent with two very different lines of research. First, writing is a constructive act. To write requires active thinking by the author to communicate emotional nuance in a way that is understood by the reader as intended, and by doing so, writing may support emotion understanding. The Production-Distribution-Comprehension account of spoken

language (MacDonald, 2013) sees establishing comprehension for the listener as central to the act of production by the speaker. Language experience flows from this production-comprehension loop, as sensitivity to the distributional properties of language builds over time and is drawn upon to support future processing encounters. This framework is attractive as it builds explicit links between language learning and language processing. Extending it to children's writing, the purposeful and productive act of selecting language to convey an intended message to the reader provides an opportunity for children to craft emotion language to capture shades of meaning (the scowls, frowns, and growls, for example). On this view, the act of writing might provide a vehicle for better understanding the emotional content of what children are writing about, and in turn, this will further tune and refine emotional understanding.

Second, a different body of work has noted that patterns of language use in expressive writing are associated with measures of well-being in adolescents, leading to the suggestion that linguistic analysis of young people's writing is a promising method for detecting mental health concerns (Cohen et al., 2022; Nook, 2023; Nook et al., 2022; Shearer et al., 2021). On this view, a piece of writing can be seen as a reflection of child's emotional understanding and their internal mind, and it is therefore reasonable to ask whether developmental changes in the use of emotional language are evident in children's own writing. While this has not been examined in detail, previous studies have found that children express and experience their emotions differently through development (e.g., Bailen et al., 2019; Thompson, 1991). Children's knowledge of emotion specific vocabulary also increases across development, with the size of the emotional lexicon doubling every 2 years between 4 and 11 years old (Baron-Cohen et al., 2010). Using a novel emotion vocabulary assessment, Nook et al., (2020) provided further evidence that knowledge of emotion specific vocabulary increases across childhood and while it plateaus at around 11 years of age, the abstractness of young people's description of an emotion continues to increase through to age 18 years. Developmental change manifests not just as an increase in the amount of emotion vocabulary size, as older children from secondary schools produce significantly more negative emotion words than younger children from primary schools (Doost et al., 1999). This runs parallel with longitudinal evidence from self-reported mood measures showing that

the frequency of negative emotion increases and positive emotion decreases from childhood into late adolescence (Larson et al., 2002; Weinstein et al., 2007).

While these two lines of research give reason to propose that children's own writing is associated with their knowledge of emotion specific vocabulary and emotional understanding, the current evidence base is severely limited, especially regarding large-scale writing samples. While future work should seek to establish the precise nature of any relationship between emotion language and emotion understanding through development, a critical first step is to capture the emotional content of children's writing, at scale and across age groups. We therefore sought to capture trends in the use of emotion-specific language across childhood, focusing on the special form of written language.

Gender differences in language and emotional development

There is evidence that gender differences in emotional development might be reflected in children's writing. In a small scale study of 8-10 year-old's responses to a topic prompt, Kanaris (1999) found that girls focused more on description and elaboration and used more diverse verbs and adjectives, whereas boys were more egocentric, using more first-person singular pronouns. O'Kearney and Dadds (2004) asked 303 adolescents (12-18 years) to describe in writing how they felt after listening to vignettes designed to promote feelings of anger or fear. Analysis of the writing samples showed an increase in the use of complex emotion terms through adolescence overall, and girls used more inner-directed terms (e.g., sad, embarrassed, guilty, ashamed) than boys.

This difference in emotion word usage in writing is consistent with the claim that boys and girls may experience and express emotions differently. Chaplin and Aldao (2013) conducted a meta-analysis on gender differences in emotional expression from infancy to adolescence, using data from behavioural observations and questionnaires. Girls showed more positive and internalising emotions, such as sympathy, while boys expressed more negative and externalising emotions, such as anger. These observations have been related to girls having greater exposure to emotional talk from parents (Shablack & Lindquist, 2019). As children get older, experience with *written* language (i.e., the

language content that children read or listen to in the context of shared reading) provides input that is gendered, as shown by the gender stereotypes that are pervasive in natural language, including books written for children (e.g., Charlesworth et al., 2021; Hsiao, et al., 2021; Lewis et al., 2022). Relating language experience back to children's own written language production, the subtle and complex linguistic choices made by children through the process of writing may reflect or further refine gender differences in social-emotional development. For example, Hsiao et al., (2021) found that children's own stories reflected the gender biases seen in children's language experience, and this patterned differently for girls and boys.

It is important to note that the existing evidence base on gender differences in emotional expression in children's writing is small; it also comes from the analysis of texts that were elicited by explicit topic prompts or questions whereby adults request an emotional response based on a hypothetical situation. Children's responses may therefore be sensitive to gender-related "display rules" and gender stereotypes (e.g., O'Kearney & Dadds, 2004). Topic prompts also tend to be based on a prototypical example, such as conflicts between close friends, and a broad emotion category, such as anger or fear. This might serve to constrain or direct children's emotion expression in writing. Our investigation was prompted by the need to better understand how emotion is expressed in writing by girls and boys in less primed contexts, using large datasets.

Large-scale language analysis

Data-driven approaches to analysing language corpora can reveal patterns in how people use language, and from this, help address psychological questions (Jackson et al., 2022, p. 807). Sentiment analysis seeks to identify attitudes and affect in natural language data (Wankhade et al., 2022). Relevant to our investigation, Hipson and Mohammad (2020) used a lexicon-based "bag of words" approach to sentiment analysis to study the emotional content of poems written by children in Grades 1 to 12 (6-18 years). This involved first obtaining sentiments of individual words from human judgement data, and then calculating a score for each poem, averaged across its words. They then analysed sentiments along three emotion dimensions (valence, arousal, dominance) and four discrete

emotion categories (anger, fear, sadness, joy). According to Warriner et al. (2013), valence refers to the pleasantness of a word and the extent of its positivity or negativity. For example, “holiday” has a positive valence whereas “prison” is negative. Arousal refers to the intensity of emotion provoked by a word with “calm” being low arousal and “active” high arousal. Dominance refers to the degree of control exerted by a word. This ranges from weak (e.g., “fatigue”) to powerful (e.g., “grand”). To capture developmental change, Hipson and Mohammad (2020) modelled these different components across ages. They found that poems written by older children showed more negative valence relative to those written by younger children, whereas arousal and dominance increased with age. Poems written by older children also showed increased intensities for all four discrete emotions. They also found that poems written by girls were more positive than poems written by boys. These findings show that sentiment analysis is a useful tool to explore children’s writing development. However, poetry is a particular genre, and other types of writing might pattern differently.

The present study

To capture the emotional content of children’s writing at scale, we adapted Hipson and Mohammad’s methods to track sentiments in a large sample of stories (N > 100,000) written by 7–13-year-olds. This enabled us to explore two hypotheses regarding age and gender. First, in keeping with the view that language is critical in the construction of emotion, we predicted that the children’s stories will show decreasing positivity with age, in line with previous research on emotional development of increased frequency of negative emotions and the increased use of negative emotion words (Doost et al., 1999; Larson et al., 2002; Weinstein et al., 2007). To date, methods have probed emotions directly. In contrast, the stories in our corpus were not written in response to a prompt or instruction (see Methods). They therefore provide an opportunity to track emotional content in more naturalistic circumstances. Our second hypothesis concerned gender differences. Following previous work in emotional development reviewed above, we predicted that stories by girls will contain more positive sentiments than those written by boys.

In summary, writing provides a vehicle for children to express emotion, and plausibly, the experience

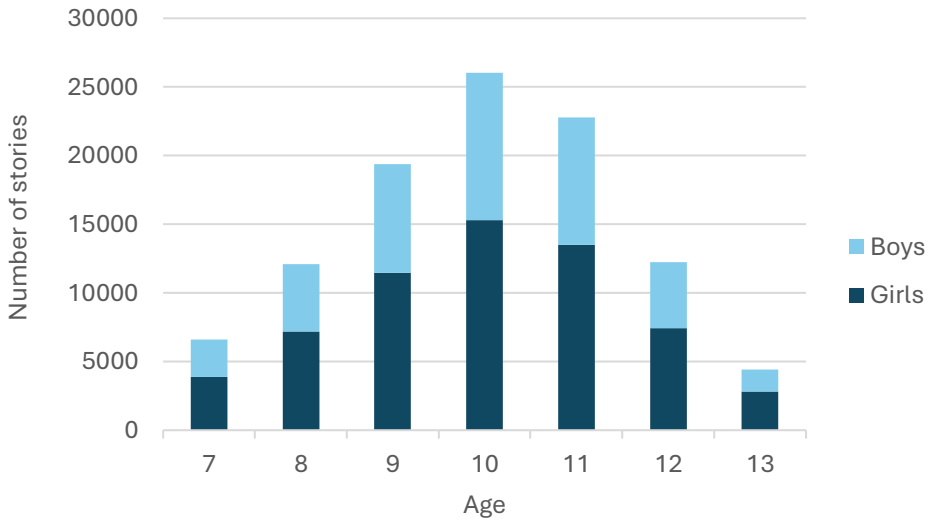
of communicating intent via writing might support the ongoing construction of emotion and language through childhood and adolescence. However, little is known about how and when children use language to communicate emotion in writing. As a first step to address this, our goal was to track sentiments in narrative writing across age and gender, using Hipson and Mohammad's (2020) methods. Note that this type of lexicon-based bag of words approach may not be good at managing the complexity of language in context. A sentence may mean different things and express different sentiments from its constituent words in isolation, more so if negation and other variations of language are considered. For example, after removing stop words such as "not", the sentence "I am happy" has the same emotional valence as "I am not happy"; clearly, this is not an accurate characterisation of the sentiment being expressed across the two sentences. With this in mind, we validated our findings using two different sentiment analysis methods that are more robust to concerns about negation.

Methods

The Oxford Children's Language Corpus

We analysed 103,564 stories written by 7- to 13-year-old children and submitted to a UK children's writing competition in 2019. BBC 500 Words was an annual national competition that invited children (aged 5-13 years until 2020, and then aged 5-11 when the competition returned in 2023) to submit a story on any topic; the only constraint was that it should be no longer than 500 words (we excluded entries from children aged 6 and under as the number of entries was relatively low). This resource provided a naturally occurring language sample not generated for assessments or any prompts or cues, allowing us to analyse children's own writing, free from constraints on time, instruction, or topic. Figure 3-1 shows the distribution of stories by age and gender. More girls (59.44%) contributed stories than boys (40.56%), and most stories came from 9- to 11-year-olds.

Figure 3-1. Number of stories by age and gender.



Note. Figure reproduced from Dong, Hsiao, et al. (2024).

We first followed Hipson and Mohammad's (2020) procedures. Tokenisation of the full corpus ($N_{stories}$ = 103,564) resulted in 46,697,930 word tokens (218,914 unique). We excluded 23 stories (3 empty files and 20 that contained only one word token; upon inspection, these contained a single letter or a random letter string). All other stories were retained for analysis ($N_{stories}$ = 103,541; $M_{words/story}$ = 451.01, $SD_{words/story}$ = 92.16).

We removed stop words (e.g., articles, prepositions) using the in-built stop word list in R (R Core Team, 2022). Emotion associations for all remaining words were extracted from the National Research Council Valence, Arousal, and Dominance (NRC VAD) lexicon v1 (Mohammad, 2018b) and the NRC Emotion Intensity (NRC EI) lexicon v0.5 (Mohammad, 2018a).[1] The NRC VAD lexicon contains about 20,000 commonly used English words that have been scored on valence, arousal, and

[1]Details of how the two lexicons were constructed can be found on this website: https://saifmohammad.com/WebPages/nrc-vad.html

dominance, with each dimension ranging from 0 to 1. The NRC EI lexicon v0.5 contains around 6,000 words from the NRC Emotion Lexicon (EmoLex) (Mohammad & Turney, 2010, 2013). These words are associated with anger, fear, sadness, or joy and each word is rated for the intensity of the associated emotion, ranging from 0 to 1.

At this stage we excluded another 49 stories as they included fewer than 5 words from the NRC VAD. This ensured that each story contained enough words for averaging in subsequent analyses. Table 3-1 shows that the number of emotion words per story increased with age. For each story, we calculated its average valence, arousal, dominance, and emotion intensity scores, based on the words it contained.

Table 3-1. Descriptive statistics of stories in different age groups

Age	Number of stories	Mean (SD) Number of emotion words per story
7	6,605	118.14 (43.70)
8	12,064	127.47 (37.94)
9	19,372	134.77 (31.90)
10	26,021	138.11 (28.90)
11	22,771	140.20 (27.10)
12	12,234	141.79 (26.34)
13	4,425	144.18 (25.73)
Total	103,492	136.12 (31.67)

Note. Table reproduced from Dong, Hsiao, et al. (2024).

Analysing developmental trends and gender differences in sentiments

We used *Generalised Additive Models (GAMs)* to analyse the children's sentiments by age as these allowed us to model a smooth relationship between the predictor variable age and the outcome variables without imposing strict parameter values on the relationship (Hastie & Tibshirani, 1990; Wood, 2017). This flexibility for capturing non-linear trends is very useful for exploratory analysis on a topic that is not well explored. Nonlinear trends in emotions over age were modelled using the *mgcv* package (Wood, 2017) in R, using Penalised Iterative Least Squares to penalise model fit as smoothing became more complex. This minimises the Generalised Cross Validation score (an index of model misfit that increases with respect to least squares and model complexity).

GAMs are additive models and do not allow for multiplicative AxB interaction terms. We investigated gender as a pseudo-interaction term, using the "by" argument: valence ~ s(Age, by = factor(Gender)). This tested for age effects at each level of gender.

Validating sentiment trends

The first validation approach used the in-built Python package *TextBlob* (Loria, 2018). This is another lexicon-based approach but takes into consideration negation, modifiers, and sentence context when determining the meaning and polarity of individual words. Polarity is the extent of positivity or negativity of the text, ranging from very negative (-1) to very positive (+1). We directly applied the *TextBlob* package without further finetuning.

Our second approach used a machine learning technique, namely a sub-model of *Bidirectional Encoder Representations from Transformers* (*BERT)* (Devlin et al., 2019). *BERT* considers the context for each occurrence of a given word to generate word embedding representations for each word. For this validation approach, we directly applied the model *"nlptown/bert-base-multilingual-uncased-sentiment"* from Hugging Face, which was pretrained on product reviews and fine-tuned for sentiment analysis. The output of the model was one of 5 sentiment classes, from negative to positive. To get a more nuanced sentiment score, we used softmax function to obtain the probability of all 5

sentiment classes by exponentiating the model's output logits and normalising them so that they summed up to 1. From the resulting matrix, we assigned a weight to each sentiment class to obtain resulting sentiment score for each story. The score was a scalar value that ranges from -1 (very negative) to 1 (very positive).

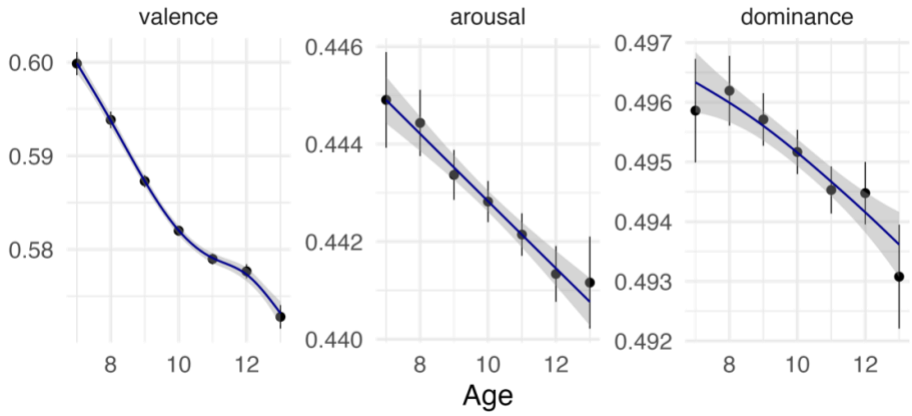
Analysis scripts are available on the Open Science Framework (http://tiny.cc/rainy_thesis_chapter3).

Results

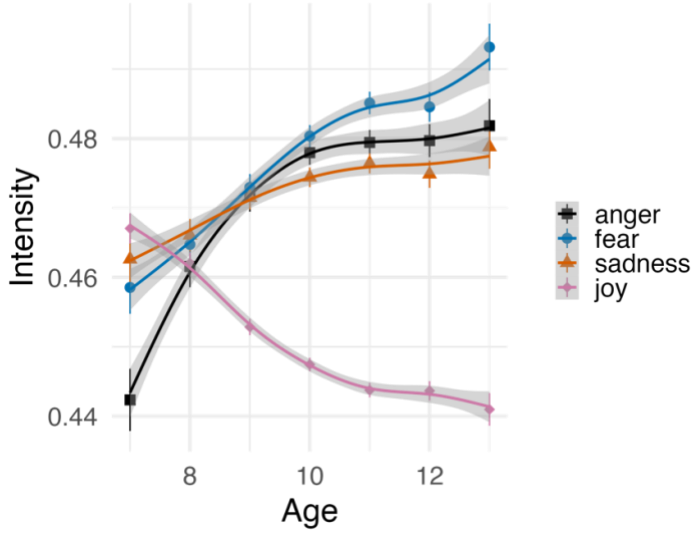
Developmental trends in sentiments

GAMs were run for each of the three dimensions (valence, arousal, dominance), and the four emotion categories (anger, fear, sadness, joy). Age was entered as a predictor in all models. We controlled for the increase in word count across ages (see Table 3-1). Figure 3-2(a) shows that mean valence, arousal, and dominance scores decreased with age (all ps < .001). Figure 3-2(b) shows the intensity of the three negative emotions (anger, fear, sadness) increased with age whereas the intensity of the positive emotion (joy) decreased (all ps < .001). Thus, across emotional dimensions and categories, children's writing showed decreasing positivity with age.

Figure 3-2. Trends in (a) valence, arousal, dominance and (b) anger, fear, sadness, and joy intensity by age.



(a) Developmental trends of valence, arousal, and dominance



(b) Developmental trends of anger, fear, sadness, and joy

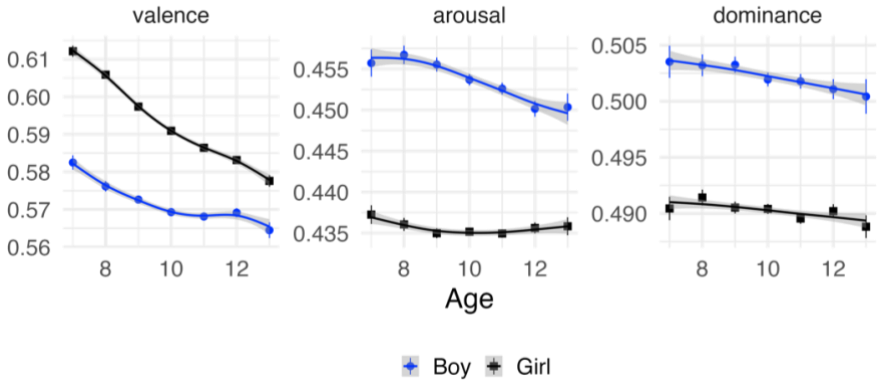
Note. The grey band represents 95% confidence intervals around the smooth fit. Figure reproduced from Dong, Hsiao, et al. (2024).

Gender differences in sentiments

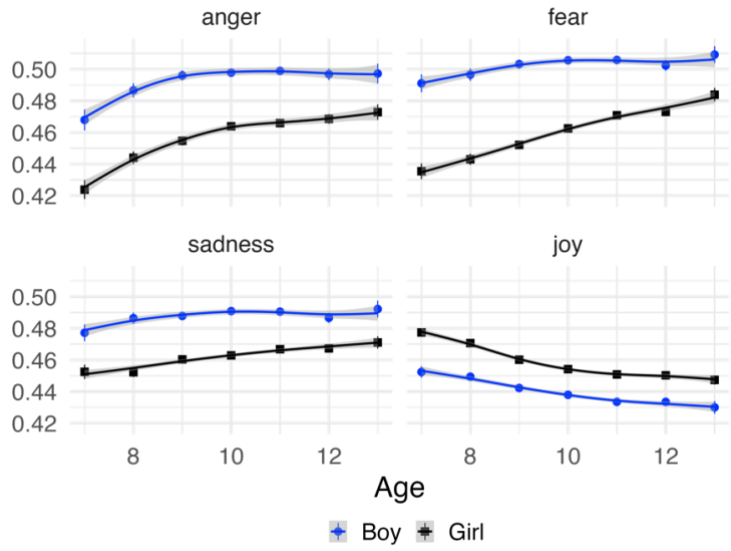
As plotted in Figure 3-3, boys and girls showed a decline in valence scores with age ($p < .001$), but stories by girls were consistently higher in valence than those written by boys. The decline in arousal was mainly due to boys, with no significant change in arousal scores across age in the stories by girls

($F < 1$, $p = 0.36$). These effects were statistically significant ($ps < .001$), with the exception of arousal scores for girls, as noted above. Stories by boys showed higher dominance scores, higher negative emotion intensity, and lower positive emotion intensity than those written by girls. Appendix 3 provide examples of usage by boys and girls.

Figure 3-3. Trends in (a) valence, arousal, dominance and (b) anger, fear, sadness, and joy intensity by age and gender.



(a) Gender difference in valence, arousal, and dominance



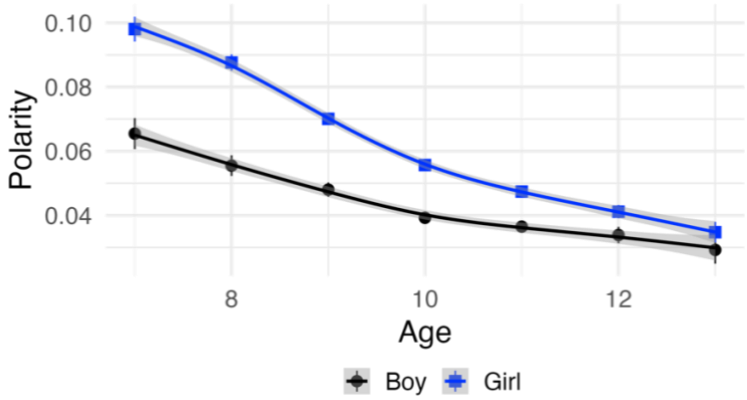
(b) Gender difference of anger, fear, sadness, and joy

Note. The grey band represents 95% confidence intervals around the smooth fit. Figure reproduced from Dong, Hsiao, et al. (2024).

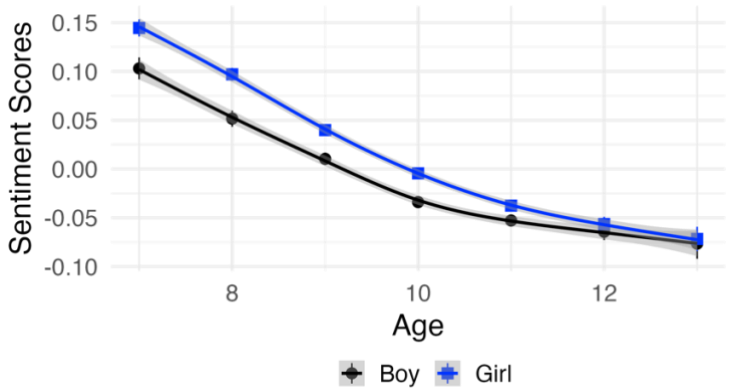
Validating sentiment trends

Figure 3-4(a) shows the trends in polarity scores of stories generated from the *TextBlob* model. As the age increased, the polarity of stories decreased. Figure 3-4(b) shows that sentiment scores generated from the *BERT* model also declined across age. *GAMs* showed these trends to be statistically significant (*p*s < .001). For both validation methods, stories by girls had more positive sentiments, especially for the younger age groups.

Figure 3-4. Trends in (a) polarity scores and (b) sentiment scores in children's story writing



(a) Trends in polarity of stories (*TextBlob*) by age and gender



(b) Trends in sentiment scores of stories (*BERT*) by age and

Note. The grey band represents 95% confidence intervals around the smooth fit. Figure reproduced from Dong, Hsiao, et al. (2024).

Finally, we computed correlations between the three measures of valence, polarity (returned by *TextBlob*), and sentiment (returned by *BERT*). As seen in Table 3-2, the three measures are positively correlated ($p < .001$).

Table 3-2. Correlations between measures of valence, polarity, and sentiment (N = 103,492)

Measures	Valence	Polarity	Sentiment
Valence	-		
Polarity	0.56	-	
Sentiment	0.28	0.28	-

Discussion

Although language is considered critical to the development of emotion, the contribution of written language has been largely ignored, and we know very little about how children construct emotion in their own writing. As a first step to understanding this, our study used a data-driven approach to capture the sentiments in children's narrative writing, sampling from a large corpus of stories written by 7- to 13-year-old children and focusing on positivity over time as well as gender differences. We first adapted and extended Hipson and Mohammad's (2020) approach and then replicated our findings using more robust sentiment analysis methods. While the three measures differed in their scoring methods and abilities to handle negation, they converged to show that the sentiments expressed by children became more negative with age. Stories written by girls contained more positive sentiments than those written by boys, especially for the younger age groups.

Decreasing positive sentiments by age

The decrease in positivity with age mirrors the pattern seen in children's poems (Hipson & Mohammad, 2020). Although broadly similar, there were some differences in the age-related change in the emotional content of stories vs poems, such as arousal, dominance, and joy intensity, which all decreased with age in children's stories, but increased with age in children's poems. These disparities might reflect inherent differences across the two genres; note also that stories were generally longer than the poems, and our age range was slightly younger too.

Why might children use more negative words as they get older? One explanation is that children experience more negative words with age. Using age-of-acquisition norms as a reference, Ponari et al. (2018) estimated that abstract words account for only 10% of vocabulary in 4-year-olds; this rises to over 40% for 12-year-olds. They also identified 8.5 years as a peak period for learning new abstract words, complementing the sharp increase in emotion knowledge seen at that time (Baron-Cohen et al., 2010). As child-directed speech contains many more positive than negative words (Ponari et al., 2018), the explosion of emotion vocabulary – and especially negative language – in mid-childhood is a likely consequence of accumulated experience with book language (Nation et al., 2022). As independent reading takes off, so too does the opportunity to learn new emotion words and use them in writing. The relative decrease in positivity with age also tracks patterns in language evolution. In English, for example, there are more negative words than positive words, and across different languages, Jackson et al. (2023) found that negative valence correlates with faster cognate replacement over time. This results in a greater variety of negative emotion words being added to the lexicon and subsequently being available for children to use, and to shape on-going emotional development, in line with language being a key ingredient in the construction of emotion (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017).

Another interpretation of the increased negativity with age is that children may choose to write about darker and more complex topics as they grow older, with the increasing understanding that negative information draws more attention (Bahn et al., 2017; Jackson et al., 2023; Pratto & John, 1991). These darker topics may invite the use of more negative words. Content analysis and topic modelling (e.g.,

T. L. Griffiths et al., 2007) could be used to explore the nature of the topics that children write about in future studies.

Alongside increased exposure to written language via reading, we speculate that learning to write and the productive act of writing provides powerful opportunities for children to construct their emotional worlds. Writing requires children to construct a language representation of the world they wish to convey into the mind of the reader. Language production shapes comprehension more generally (McDonald, 2013), and there is some evidence that writing drives the development of reading comprehension (Zagata et al., 2023). Emotion production may drive emotion comprehension, and vice versa. While children's stories provide a window to understand how they represent and describe the emotions of characters, it is important to note that the fictional world they create may not be representative of their own emotions. Although the decreasing positive sentiments in children's writing run parallel with their decreasing self-reported moods across development (e.g., Larson et al., 2002), our study is not able to establish a direct link between the two. It would be interesting to directly compare the linguistic content of children's writing when they are composing a story vs. reflecting on their own feelings and experiences, especially given the interest in writing as a method to identify mental health concerns in adolescence (Cohen et al., 2022; Shearer et al., 2021). Experimental work is needed to establish whether there are direct relationships between children's experience with written language, their writing, and their own emotions.

Gender difference in sentiments in children's writing

In line with our prediction, stories by girls were more positive than those by boys, especially for the younger age group. This extends previous observations from elicited tasks and self-report to narrative writing. Gender differences have been associated with differences in child-directed speech, with caregivers using more emotional language with girls (for review, Shablack & Lindquist, 2019). Adding to this, book language itself is also associated with systematic gender differences (Lewis & Lupyan, 2020). For example, Lewis et al. (2022) measured word-gender association in a children's book corpus and asked adults to rate the femininity or masculinity of words. They found that words

rated as more feminine or more masculine were also more likely to be associated with characters of the respective gender. There is also evidence of androcentrism in children's books, and of this being reflected in children's own writing (Hsiao et al., 2021). Gender differences may be perpetuated in that girls are more likely to read books featuring girl characters and girl-related content, and vice versa for boys (Charlesworth et al., 2021; Lewis et al., 2022), and thus their own word choices reflect the language they experience in stories they read. We suggest that children's writing may serve as a valuable resource to understand emotional development in boys and girls (Chaplin & Aldao, 2013) and provide a foundation for research that seeks to identify potential young people's mental health concerns from analysis of their writing (Cohen et al., 2022; Shearer et al., 2021).

Limitations and future direction

While converging evidence demonstrated a decrease in positive sentiments in children's writing, some limitations need to be acknowledged. Using a naturally occurring dataset allowed us to examine real-world language use at scale, but we had no control over how the dataset was constructed, and metainformation is limited. For instance, the cross-sectional datasets obtained from a national competition might not fully represent the population. The competitive nature of the submission likely attracted more children with stronger language skills, introducing potential selection bias and underrepresenting those with weaker language abilities. Also, parental input into the writing was unmonitored and could further compromise the representativeness of this sample. Nonetheless, we note that the competition was open to all and was advertised extensively on local and national media and via schools. Entries were received from all parts of the UK, spanning rural and urban neighbourhoods. Unfortunately, however, we were unable to access detailed demographic information for all entries.

Moreover, we note that using pre-trained models without fine-tuning to children's stories may not fully capture the actual sentiment (Singh et al., 2017), although Werlen et al. (2019) argued that applying machine learning algorithms to different contexts was still effective. To validate the observed developmental trends in sentiments, future studies should employ longitudinal datasets and

incorporate some hand-coded data for fine-tuning the language model. Note that we did not calculate effect size as it is not clear how this should be done or interpreted when using Generalised Additive Models. Given the potential issue of overpowering given corpus size (Egbert et al., 2022), however, it is likely that our effect sizes are small and this is perhaps not surprising given the task was open-ended, without topic prompts or specific instructions. More generally, there is a pressing need for experimental work on the connections between children's writing and emotional development. Such work will enable causal inference and help us understand the theoretical and practical significance of small effects; this is especially important if linguistic analysis is to be applied to mental health and clinical concerns (e.g., Cohen et al., 2022).

Many questions remain regarding the interplay between language and emotion and how this unfolds over development (e.g., Lindquist, 2021). We propose that experience with written language provides children with access to particular forms of emotion language that is then reflected in how children shape their own writing to communicate emotional content. Our study demonstrates the utility of large-scale data driven approaches, but our findings need to be complemented by future work taking a more experimental approach. These will reveal whether and how experience with written language interacts with the acquisition of emotional knowledge and vice versa, and how this becomes reflected in children's own writing.

Having looked at emotion words in language written for and by children in Chapters 2 and 3, I now shift focus to see how the emotion provided by the narrative context might support word learning in both adults (Chapter 4) and children (Chapter 5).

Chapter 4. Adult Word Learning in Emotional Context

This chapter is based on the following paper (in press):

Dong, Y., Mak, M. H., Hepach, R., & Nation, K. (2024). EXPRESS: Learning New Words via Reading: The Influence of Emotional Narrative Context on Learning Novel Adjectives. *Quarterly Journal of Experimental Psychology*, 0(0). https://doi.org/10.1177/17470218241308221

Chapter summary:

People learn new words in narrative contexts, but little is known about how the emotional valence of the narrative influences word learning. In a pre-registered experiment, seventy-six English-speaking adults read 30 novel adjectives embedded in 60 short narratives (20 positive, 20 negative, and 20 neutral valence). Both immediately after and 24 hours later, participants completed a series of post-tests including speeded recognition, sentence completion, meaning generation, and valence judgment. Results showed that participants learned both the novel word form and its meaning. Compared to novel words experienced in the neutral contexts, those read in the emotional contexts (both positive and negative) showed better learning of orthographic form in the immediate post-test, but only those read in the negative context were recognised with greater accuracy in the delayed post-test. Furthermore, the valence of the context was reflected in the word meanings participants generated for each novel word, suggesting that word valence can be inferred from the valence of the contexts. Results from sentence completion and valence judgment were mixed, depending on the task demands. These findings are discussed with reference to theories of affective embodiment and the implications for learning abstract words are considered.

Introduction

In language, emotional valence refers to the pleasantness of a word and the extent of its positivity or negativity (Warriner et al., 2013). This broadens the definition of emotion words from just describing

an emotional state, such as *happy* or *sad*, to all content words. Word valence influences how early and how well a word is learnt (e.g., Kousta et al., 2011; Ponari et al., 2020); it also influences lexical processing in adults (e.g., Vinson et al., 2014). Most existing research relies on valence norms in which people rate the positivity of familiar words (e.g., Warriner et al., 2013). For an unfamiliar novel word, however, how do people learn its valence? One possibility is from the valence of the context in which it appears. In line with this, Snefjella and Kuperman (2016) reported a positive correlation between word valence and contextual valence, defined as the aggregate valence of the five content words immediately before and immediately after the word in text samples taken from a large corpus of email newsgroup postings. Word-learning experiments (e.g., Snefjella et al., 2020) have also investigated how valence of the word might be learned from emotional contexts, but the results have been mixed. In this pre-registered study, we investigated whether and how variations in the valence of emotional narrative context might influence the learning of new adjectives.

Word valence is known to influence lexical processing. In lexical decision for example, emotional words, whether positive or negative, are processed faster and with greater accuracy than neutral words, a phenomenon that persists regardless of the mode of word presentation (Kousta et al., 2009; Ponari et al., 2018; Scott et al., 2009, 2012). Further support comes from neuroimaging and electrophysiological studies, which demonstrate that words with more extreme valence elicit distinct neural responses compared to neutral words (Pauligk et al., 2019; Vigliocco et al., 2014; Yao et al., 2016). This processing advantage extends beyond isolated words to those presented within sentences, affecting both lexical processing and memory (Bayer et al., 2010; Scott et al., 2012).

Emotional valence predicts age-of-acquisition ratings and emotionally valenced abstract words tend to be earlier acquired than neutral ones (Kousta et al., 2011). Ponari et al. (2018) suggested that emotional valence provides a bootstrapping mechanism for acquiring abstract concepts. However, this valence effect is not uniformly observed across all age groups but is particularly pronounced in children aged 8 to 9 years (Lund et al., 2019; Ponari et al., 2018). Moreover, valence also impacts children’s learning and memory for newly taught abstract words, with emotionally valenced words

being learned better and defined more accurately than neutral words (J. M. Kim et al., 2020; Ponari et al., 2020). These findings align with the affective embodiment account which proposes that emotional content aids in grounding abstract word meanings in emotional experiences, providing a motivational relevance that heightens processing efficiency (Vigliocco et al., 2014). However, while this account suggests that emotional valence facilitates the acquisition of abstract words, it does not have clear predictions about its directionality, i.e., whether positive valence or negative valence better supports word learning.

There are mixed findings as to the directionality of any valence influence on word processing and learning. Some studies have found a positivity advantage (e.g., Kuperman et al., 2014; Ponari et al., 2018; Yao et al., 2016), which can be explained by the Pollyanna principle, a global tendency for humans to remember pleasant things more accurately (Matlin & Stang, 1978). Unkelbach et al. (2008) proposed the informational density hypothesis, where positive information, being more elaborated and densely clustered, is processed faster than negative information. The greater interconnectivity of positive words in a denser semantic network might result in it being activated faster during word processing. Other studies, however, have also found a negativity advantage (e.g., Estes & Verges, 2008). This has been explained in terms of an increased vigilance for negative stimuli (Pratto & John, 1991) by which individuals have an intrinsic tendency to focus attention on negative stimuli. This can lead to more in-depth processing of negative information and therefore better recognition memory (Ortony et al., 1983). Similarly, the Negative Emotional Valence Enhances Recapitulation model (Bowen et al., 2018) further highlighted the role of negative valence in enhancing the reactivation of sensory details over time, suggesting that already in early ontogeny negative events and stimuli may be encoded and retrieved with greater sensory fidelity (Vaish et al., 2008). While both positive and negative advantages have been reported, it is difficult to draw clear conclusions regarding directionality, not least because different studies use different methods and might therefore tap into different processes.

Word valence is usually determined from large-scale rating studies where participants rate the valence

of individual words on a scale (e.g., Warriner et al., 2013). Affective ratings can be predicted from contextual variables such as contextual diversity (Recchia & Louwerse, 2015), and Kuhlmann et al. (2017) found that the valence of semantic neighbours within a word's associative network influences the perceived valence of neutral words. This led them to consider valence as a "semantic super-feature". In line with this, Snefjella and Kuperman (2016) reported a positive correlation between word valence and contextual valence. This suggests that words acquire a valence that reflects the overall emotional tone of the contexts in which they are used. They also found that the contextual valence of a word predicts lexical decision performance, even when the influence of word valence was controlled, reinforcing the idea that a word's contextual history influences lexical processing (Hsiao et al., 2020). From this background, experiments have investigated the influence of emotional context on the learning of new words. Snefjella et al. (2020) asked native speakers of English to learn nine novel nouns (e.g., *plurk* as a substitute for "the real word"), each embedded in five short passages that were designed to be neutral, negative, or positive. In the subsequent post-tests, participants showed clear evidence of word-form and meaning learning, with the positive condition resulting in the best learning outcomes. Importantly, all the novel words acquired emotional connotations, suggesting that there was transfer of valence from contexts. Using the same paradigm, Lana and Kuperman (2023) investigated the learning of novel words that denoted either abstract (e.g., religion) or concrete nouns (e.g., tool). In line with Snefjella et al. (2020), they found that positive contexts boosted word learning overall, but contextual transfer of valence was only evident for concrete, but not abstract nouns. While other experiments have investigated word learning while varying emotional context, findings are difficult to compare as different methods have been used. For example, Driver (2021) found better word learning when novel words denoted neutral concrete meanings and when novel words were embedded in neutral or negative emotion-laden texts, yet Frances et al. (2020) reported a facilitative effect for novel concrete words learned in positive contexts in relation to neutral contexts; note however they did not include any negative contexts. Taken together, these studies consistently show that a novel concrete word can acquire valence from the emotional tone of its surrounding text. Less clear, however, is whether this generalises to abstract words and whether positive or negative contexts (or both) support word learning, and why.

Existing studies have focused predominantly on nouns. Plausibly, words from other grammatical classes, such as adjectives, might show a different pattern. Compared to nouns, adjectives tend to be more abstract, and their meanings might be more context-dependent (Davies et al., 2020; Dawson et al., 2021). It is thus likely that the effect of contextual valence may be more important for this word class. We therefore focused on novel adjectives and used a naturalistic reading procedure to investigate the effect of contextual valence on word learning. We asked whether people learn the valence of novel adjectives from positive, neutral, and negative contexts immediately after reading. We also asked how well people remembered the newly learned words 24 hours later. This allowed us to investigate across the two test points the hypothesis that emotional narrative context facilitates novel word learning through affective embodiment. We have the following three predictions:

1. Participants would learn novel word forms from reading short narratives, especially in more emotional (positive and negative) contexts.

2. Participants would infer novel word meanings from reading short narratives, especially in more emotional (positive and negative) contexts.

3. Participants would infer the valence of novel words from the linguistic context in which they appear.

Method

Participants

Eighty-seven participants (42 females, 45 males) were recruited through Prolific and completed both sessions remotely. Their ages ranged from 18 to 30 years old (M_{age} = 25.72, SD_{age} = 3.27). All participants reported to be native English speakers based in the UK, have normal or corrected-to-normal vision, and no history of dyslexia or other language difficulties. They all provided consent before taking part. Following our pre-registered exclusion criterion, 11 participants were excluded from all the analyses due to them failing more than 20% of the attention checks. The final sample size

was 76 (34 Females, 42 Males; range 19 to 30 years, M_{age} = 25.58, SD_{age} = 3.30). They received a £7.5 payment for their participation via Prolific.

Design

There was one independent variable, contextual valence, with three levels: neutral, negative, and positive. This was manipulated within participants. Accuracy and reaction time (RT) were measured and served as dependent variables. The study spanned two sessions. Session 1 consisted of a reading phase and an immediate test phase. Session 2 consisted of a delayed test phase only and was available 24 hours after participants completed Session 1. The study, including the sample size, exclusion criteria and confirmatory analysis plan, was pre-registered ahead of data collection (https://osf.io/sc4ze).

Materials

We created 60 naturalistic narratives ($M_{word\ count}$ = 18.95, $SD_{word\ count}$ = 2.65) of either positive, neutral, or negative valence (20 narratives in each condition). The sentiment of each narrative was estimated using a *Bidirectional Encoder Representations from Transformers (BERT)* model, a transformer-based machine learning technique for natural language processing (Devlin et al., 2019). *BERT* considers the entire context of words in a sentence, rather than one word at a time, allowing it to capture emotional nuances in the narratives (see Appendix 4 for more details of the approach). These narratives were also rated by a group of 19 native English speakers, who did not participate in the main study, on a Likert scale of 1-7, how positive/negative each narrative makes them feel. Pearson correlation shows that the valence ratings by the 19 participants correlated strongly and positively with the *BERT* ratings, $r = 0.69$, $p < .001$. Both approaches showed that contextual valence differed significantly across the three valence conditions based on one-way Analysis of Variance, with narratives in the positive condition showing the most positive valence, followed by neutral, and then by negative. Narratives across conditions were matched for their overall word count (M = 18.95, SD = 2.68) and mean length of utterance. A sample set of narratives is shown in Table 4-1 (see Appendix 4 for all

narratives and computations).

There were 70 novel words, 30 of which were target novel words and the others were foils. Each of the 30 target novel words was read by participants in two narratives of the same valence. Within each narrative, the target word appeared twice (see Table 4-1). The novel words were 6 or 7 letters long, *M* = 6.63 letters, *SD* = 0.46. They do not have a base meaning and were created to have a nonword stem plus an adjective suffix. For example, the nonword stem "thut" and the adjective suffix "ive" led to the novel word "thutive". We chose 10 adjective suffixes from a list of suffixes with high diagnosticity values for adjectives, as calculated by Ulicheva et al. (2020). The target novel words had no replacement orthographic neighbours, according to *NWatch* (Davis, 2005). Assignment of target novel words to the contextual valence condition was counterbalanced such that a novel word that appeared in the positive context for one participant appeared in the neutral or negative context for other participants. The foils were the same across participants.

Table 4-1. Example narratives in each contextual valence condition.

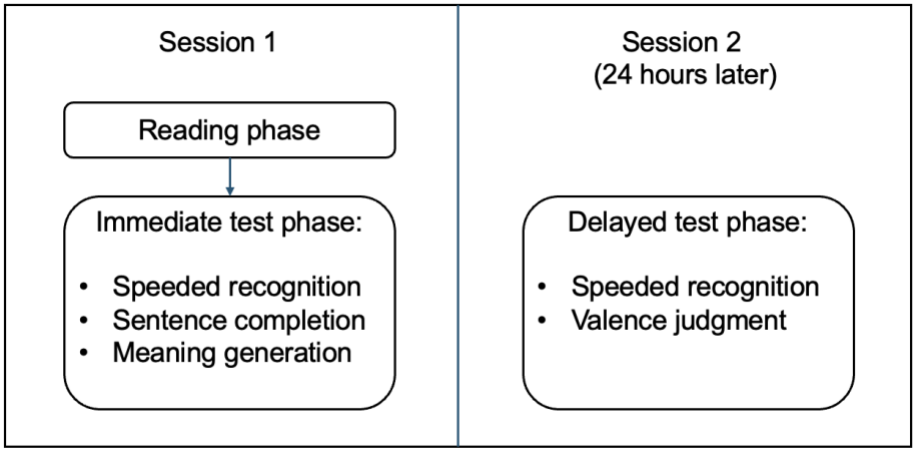
Contextual Valence	Examples
Neutral	This sopable machine was newly produced by the company. It has a sopable cover and four wheels.
Negative	I had a plarous argument with my friend today. We could not agree. Their plarous words hurt my feelings.
Positive	I am having a picial time with my family in this beautiful weather. We enjoyed the picial scenery.

Procedure

Participants signed up to the two-session study on Prolific. After reviewing the participant information sheet and providing consent, participants reported basic demographic data. The experiment was programmed and hosted on Gorilla Experiment Builder (www.gorilla.sc, Anwyl-Irvine et al., 2022).

The procedure of the experiment is shown in Figure 4-1.

Figure 4-1. Procedure of the two-session study.



(i) Reading phase. This was structured around the premise of an alien attempting to learn English who occasionally replaced English words with words from its own language when writing a diary. These served as the novel words for the purposes of the word learning experiment. Participants read the novel words in neutral, negative, or positive narratives, and each participant experienced 30 novel words embedded in 60 short narratives, evenly split into two blocks of 30. They were told to pay attention to these novel words. Each block also contained five narratives that served as attention checks, where participants answered a comprehension question (in a multiple-choice format) based on the narrative they just read. Participants who failed 20% of these checks ($N = 11$) were excluded from all analyses.

(ii) Immediate test phase. Immediately after reading the narratives, word learning was assessed via three outcome measures administered in a fixed order: speeded recognition, sentence completion, and meaning generation. For *speeded recognition*, modelled after lexical decision, participants identified whether they had previously seen a presented letter string. Each trial began with a fixation cross

displayed for 250 ms, followed by the letter strings, during which the participants pressed buttons on the keyboard to make a judgment. The next trial began as soon as a response was recorded. Accuracy and RT were recorded. Each participant responded to 60 items in total (30 target novel words and 30 distractors), presented in a random order in a single block.

In *sentence completion*, each trial showed a sentence with a missing word, and participants were asked to select the best completion for the sentence from a choice of three novel words they had encountered in the reading phase (one from each valence condition). There were 30 trials in total, 10 in each of the valence conditions. Each trial began with a fixation cross displayed for 250 ms, followed by the sentence and the three options displayed below. Participants made a judgment by clicking on the most suitable novel word option. The next trial began as soon as a response was recorded. Accuracy and RT were recorded, and no feedback was provided. The 30 trials were presented in a random order in a single block.

The final task was *meaning generation*. In each trial, participants were shown one of the 30 newly learnt words (presented in a randomised order), and they were required to type in an English word they considered to correspond with the meaning of the novel word. The words produced were cross-referenced with norms of valence for English lemmas (Warriner et al., 2013), and these values were used to assign a valence score to each response. The outcome variable is the valence of the produced words (ranging from 1 to 9, with 1 being very negative and 9 being very positive), rather than a more fine-grained meaning.

At the end of Session 1, participants completed a brief questionnaire soliciting their perceptions of the experiment, their reading strategies, and any additional comments (see Appendix 4). They were reminded that there would be a delayed test available 24 hours later, but they were not told what tests would be administered. The reading phase and the immediate test phase in Session 1 took around 30 minutes to complete.

(iii) Delayed test phase. Twenty-four hours later, participants were sent reminders, and they completed

an identical version of *speeded recognition* to Session 1, with trials appearing in random order. They then completed a *valence judgment* task (which was not administered in Session 1) in which they made a categorical judgment to indicate whether a novel word (experienced in the reading phase) was positive, neutral, or negative. This task mentioned the concept of valence explicitly; hence it was only administered at the end so it would not influence participants' performance on other tasks. Each word was displayed for 300 ms before the three options appeared. Participants made a judgment by clicking on the most suitable option. The next trial started as soon as a response was recorded. Accuracy and RT were recorded. There were 30 trials in total, presented in a random order in a single block. The delayed test phase in Session 2 required 10 minutes to complete.

Statistical analyses

As pre-registered, we fitted mixed-effect models with random effects for participants and stimuli, using the *lme4* package (Bates et al., 2015) in R (R Core Team, 2022). RT data for the correct trials were transformed to give a more normal residual distribution, based on the suggestion of the Box-Cox procedure (Box & Cox, 1964) and inspection of the qqplot (Millard, 2013). For all full models DV ~ ContextualValence + (1+ ContextualValence | participant) + (1+ ContextualValence | item) that failed to converge, we followed (Mak, Curtis, et al., 2023; Mak, O'Hagan, et al., 2023) by simplifying the random-effect structure using the R package *buildmer* (Voeten, 2023). Unless otherwise specified, after simplification, a binary logistic mixed-effects model was adopted for the dependent variable accuracy (1 or 0) and included the fixed effect of contextual valence and by-participant and by-stimuli random intercepts. A linear mixed model was fitted to the transformed RTs, with contextual valence as the sole fixed effect, and random intercepts for participants and stimuli. A likelihood ratio test was used to compare the full model to the reduced model to assess whether including the fixed factor ContextualValence significantly improved model fit.

The fixed effect was contextual valence (neutral, negative, or positive). This was dummy-coded, and the neutral condition served as the reference level, yielding two comparisons: neutral vs. positive and neutral vs. negative. Models were fitted using maximum likelihood estimates. Following these

confirmatory analyses, we also conducted an exploratory analysis comprising a direct comparison between the positive and negative conditions. Data and analysis scripts are accessible on the OSF: http://tiny.cc/rainy_thesis_chapter4and5.

Results

Mean performance in each post-test is summarised in Table 4-2. We first present the confirmatory analyses based on the hypotheses, and then the exploratory analyses.

Table 4-2. Descriptive statistics (M and SD) of performance in immediate and delayed post-tests.

Task	Speeded Recognition (Immediate)		Sentence Completion (Immediate)		Meaning Generation (Immediate)	Speeded Recognition (Delayed)		Valence Judgment (Delayed)	
Measured Variable	Accuracy	RT	Accuracy	RT	Valence score	Accuracy	RT	Accuracy	RT
Neutral	0.71 (0.18)	853 (276)	0.49 (0.18)	5078 (3027)	5.62 (0.63)	0.78 (0.15)	831 (341)	0.44 (0.20)	1409 (1019)
Negative	0.79 (0.15)	835 (203)	0.53 (0.17)	4207 (1710)	4.89 (0.73)	0.82 (0.15)	826 (315)	0.37 (0.17)	1371 (719)
Positive	0.76 (0.17)	835 (230)	0.48 (0.21)	4786 (2924)	5.98 (0.69)	0.80 (0.16)	832 (346)	0.45 (0.17)	1291 (970)

Note. RT is measured in milliseconds (ms); Accuracy is a probability from 0 to 1; Valence score is a continuous value between 1 and 9 based on Warriner et al. (2013), with larger values indicating greater positivity, and smaller values indicating greater negativity. The chance level is 0.5 for Speeded Recognition, and 0.33 for Sentence Completion and Valence Judgment,

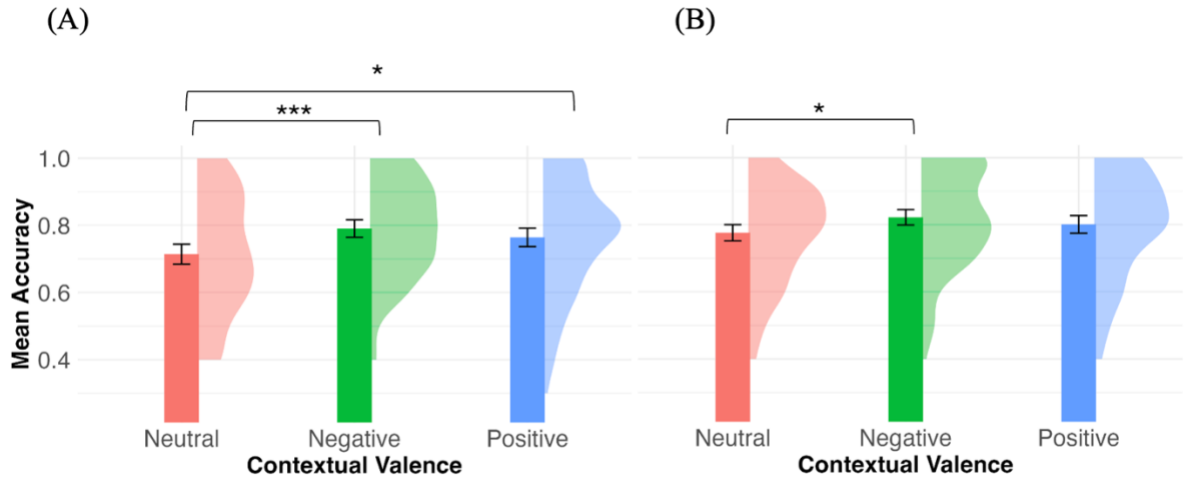
Confirmatory analysis

Word form learning assessed by speeded recognition

Word form learning was assessed via *speeded recognition* in both sessions. Starting with the immediate post-test data, we first computed sensitivity from raw responses to the speeded recognition task (number of hits, false alarms, correct rejection, and misses) using the *dprime* function in the *psycho* R package (Makowski, 2018). A one-sample *t*-test provided clear evidence that participants could distinguish learned items from distractors, $M_{d'} = 1.49$, 95% CI = [1.35, 1.63], $t(75) = 21.30$, $p < .001$.

We then compared novel words learned across the different valence conditions. Figure 4-2(A) shows the mean recognition accuracy by contextual valence in the immediate post-test. From the likelihood ratio test, contextual valence was a significant predictor for recognition accuracy, $\chi^2(2) = 12.57$, $p = .002$. As compared to words in the neutral context ($M = 0.71$, $SD = 0.18$), participants were more accurate in recognising words experienced in the negative context ($M = 0.79$, $SD = 0.15$) and the positive context ($M = 0.76$, $SD = 0.17$; negative vs. neutral: $\beta = 0.44$, $SE = 0.13$, $z = 3.50$, $p < .001$; positive vs. neutral: $\beta = 0.28$, $SE = 0.12$, $z = 2.26$, $p = .02$).

Figure 4-2. Mean recognition accuracy by contextual valence in immediate post-test (A) and delayed post-test (B).



Note. The density plots represent the distribution of the mean accuracy. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

We also tested for the differences in RT in the three valence conditions at immediate *speeded recognition*. Of all "hit" trials ($N = 1725$), 24 trials (1.3%) with RTs that were >3 SDs away from the mean RT of that participant were removed. The remaining RT data were inversely transformed to give a more normal residual distribution. There was no significant effect of contextual valence on RT, $\chi^2(2) = 0.87$, $p = .65$. Compared to words learned in the neutral context ($M = 853$ ms, $SD = 276$), there were not significant differences between novel words learned in the negative context ($M = 835$ ms, $SD = 203$) or the positive context ($M = 835$ ms, $SD = 230$; negative vs. neutral: $\beta = -0.013$, $SE = 0.016$, $t = 0.80$, $p = .42$); positive vs. neutral: $\beta = -0.013$, $SE = 0.016$, $t = 0.82$, $p = .41$).

Mirroring the findings from the immediate recognition test, there was evidence that participants could distinguish learned items from distractors in the delayed recognition test 24 hours after learning, $M_{d'} = 1.63$, 95% CI = [1.50, 1.75], $t(75) = 25.5$, $p < .001$. Accuracy data are shown in Figure 4-2(B). The fixed factor contextual valence was marginally significant, $\chi^2(2) = 5.81$, $p = .05$. Compared to words in the neutral context ($M = 0.78$, $SD = 0.15$), participants were more accurate in recognising words experienced in the negative context ($M = 0.82$, $SD = 0.15$), $\beta = 0.33$, $SE = 0.14$, $z = 2.43$, $p = .02$. There was no difference between the positive ($M = 0.80$, $SD = 0.16$) and neutral condition, $\beta = 0.18$, $SE = 0.13$, $z = 1.37$, $p = .17$.

Of all hit trials ($N = 1824$), 39 (2.1%) had an RT > 3 SDs away from the participant's mean RT and were hence removed. The fixed effect of contextual valence was not significant, $\chi^2(2) = 0.86$, $p = .65$. Compared to words appearing in neutral valence paragraphs ($M = 831$ ms, $SD = 341$), there were no significant differences between RT towards novel words in the negative context ($M = 826$ ms, $SD = 315$) or the positive context ($M = 832$ ms, $SD = 346$; negative vs. neutral: $\beta = -0.015$, $SE = 0.016$, $t = 0.92$, $p = .36$); positive vs. neutral: $\beta = -0.0053$, $SE = 0.016$, $t = 0.33$, $p = .74$).

In summary, the results of *speeded recognition* supported our first prediction that participants can learn novel word forms from reading short narratives, as indexed by the above-chance sensitivity in distinguishing learned items from distractors in both sessions. Participants were more accurate in emotional (positive and negative) contexts in the immediate post-test, and more accurate only in the negative context in the delayed post-test. There were not significant contextual valence effects in RTs.

Word meaning learning assessed by sentence completion and valence judgment

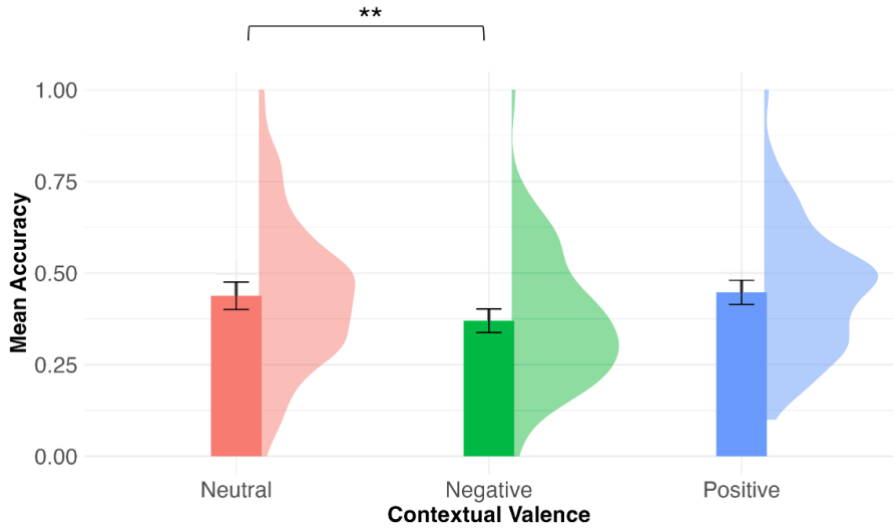
Word meaning learning was assessed with *sentence completion* in the immediate post-test and *valence judgment* in the delayed post-test. To test learning of novel word meanings in *sentence completion*, we first compared a participant's response accuracy with chance performance, which was 0.33. Using a one-sample *t*-test, there was clear evidence that participants' performance was above chance, $M_{accuracy}$ = 0.50, 95% CI = [0.47, 0.53], $t(75) = 9.86$, $p < .001$. The likelihood ratio test shows that contextual valence was not a significant predictor of sentence completion accuracy, $\chi^2(2) = 1.24$, $p = .54$. As compared to the neutral context ($M = 0.49$, $SD = 0.18$), there was no difference in accuracy for the negative context ($M = 0.53$, $SD = 0.17$), or the positive context ($M = 0.48$, $SD = 0.21$; negative vs. neutral: $\beta = 0.14$, $SE = 0.17$, $z = 0.84$, $p = .40$; positive vs. neutral: $\beta = -0.04$, $SE = 0.17$, $z = -0.24$, $p = .81$).

Of all the correct trials ($N = 1140$), those more than 3 SDs from the mean RT for that participant were removed ($N = 9$, 0.8%). The likelihood ratio test shows that the fixed effect of valence was a significant predictor of log-transformed sentence completion time, $\chi^2(2) = 7.19$, $p = .03$. Compared to reaction times for neutral sentences ($M = 5078$ ms, $SD = 3027$), participants responded faster in negative ($M = 4207$ ms, $SD = 1710$), but not positive contexts ($M = 4786$ ms, $SD = 2924$; negative vs. neutral: $\beta = -0.13$, $SE = 0.04$, $t = -2.88$, $p = .004$); positive vs. neutral: $\beta = -0.07$, $SE = 0.04$, $t = -1.52$, $p = .13$).

As above, a one-sample *t*-test showed that participants' performance was above chance (0.33) in *valence judgment* in the delayed post-test, $M_{accuracy} = 0.42$, 95% CI = [0.39, 0.45], $t(75) = 6.40$, p

< .001. Figure 4-3 shows the mean valence judgment accuracy by contextual valence. Contextual valence was a significant predictor of valence judgment accuracy, $\chi^2(2) = 11.70$, $p = .003$. As compared to novel words learned in the neutral contexts ($M = 0.44$, $SD = 0.20$), novel words in the negative conditions were less accurately judged as negative ($M = 0.37$, $SD = 0.17$), $\beta = -0.30$, $SE = 0.11$, $z = -2.77$, $p = .006$. There was no significant difference between the neutral and positive condition ($M = 0.45$, $SD = 0.17$), $\beta = 0.04$, $SE = 0.11$, $z = 0.36$, $p = .72$.

Figure 4-3. Mean valence judgment accuracy by contextual valence in the delayed post-test.



Note. The density plots represent the distribution of the mean accuracy. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Of all 954 correct trials, those more than 3 SDs away from the mean RT for that participant were removed ($N = 8$, 0.8%). A linear mixed-effects model was used with inverse square root transformed RT as the dependent variable and included the fixed effect of contextual valence and random intercepts and slope for participants. Contextual valence was a significant predictor for RT, $\chi^2(2) = 10.02$, $p = .007$. Yet, compared to words experienced in the neutral context ($M = 1409$ ms, $SD =$

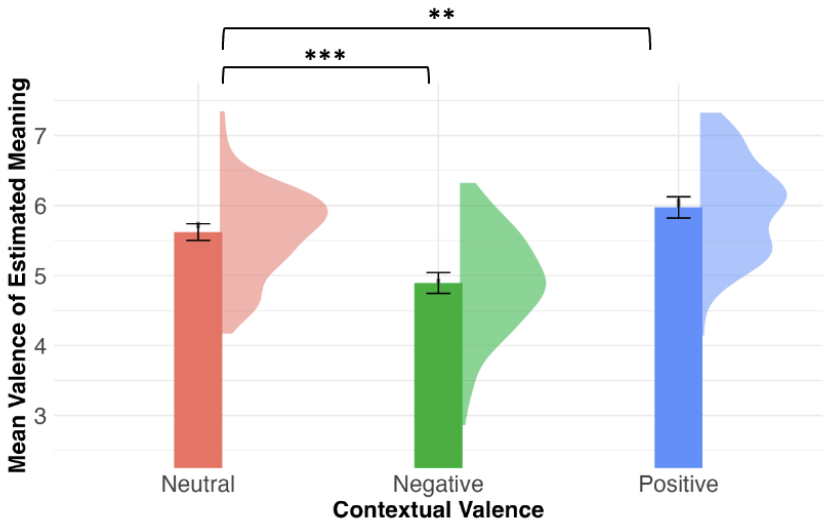
1019), there was no difference in RT for words learned in the negative context (M = 1371 ms, SD = 719), or positive contexts (M = 1291 ms, SD = 970; negative vs. neutral: β = 0.69, SE = 0.68, t = 1.03, p = .30; positive vs. neutral: β = -1.11, SE = 0.68, t = -1.62, p = .10).

These results partially support the second prediction that participants can infer novel word meanings from reading short narratives. Performance was above chance in both *sentence completion* and *valence judgment*. However, participants were not consistently more accurate or faster in the emotional contexts. We return to discuss this finding later.

Word valence inference assessed by meaning generation

Most responses (1934 out of 2280; 84.8%) had valence scores listed in Warriner et al.'s (2013) norms. Reasons for the absence of associated valence included random letter strings, "?", more than one word, or the response word was not normed. Some examples of generated words included *enjoyable, happy, terrible, good, bad, annoyed, clear, loud*. Majority of generated words are adjectives. A complete list of participant's production can be accessed through the project's OSF page. Figure 4-4 shows the mean valence score of generated meaning per participant by contextual valence. We built a linear mixed model with the valence score as the dependent variable and included the fixed effect of contextual valence and a by-item random slope and random intercepts. The likelihood ratio test shows that contextual valence was a significant predictor of estimated valence scores, $\chi^2(2) = 33.99$, $p < .001$. Compared to novel words appearing in the neutral context (M = 5.62, SD = 0.63), participants assigned more negative meanings to novel words experienced in the negative context (M = 4.89, SD = 0.73), and more positive meanings to novel words experienced in the positive context (M = 5.98, SD = 0.69; negative vs. neutral: β = -0.72, SE = 0.15, t = -4.66, $p < .001$; positive vs. neutral: β = 0.36, SE = 0.12, t = 2.98, p = .003).

Figure 4-4. Mean valence scores of generated meanings (immediate post-test) per participant by contextual valence.



Note. The density plots represent the distribution of the mean estimated valence. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Interim summary of confirmatory analyses

The hypothesis that emotional valence facilitates novel word learning through affective embodiment was largely supported by our results. The results of *speeded recognition* in both immediate and delayed post-tests supported the first prediction that participants can learn novel word forms from reading short narratives. Participants were more accurate in emotional (positive and negative) contexts in the immediate post-test in Session 1, and more accurate only in the negative context in the delayed post-test in Session 2. The second prediction that participants can infer novel word meanings from reading short narratives was partially supported by the above-chance performance of *sentence completion* and *valence judgment*. Compared to neutral conditions, there was no difference for positive conditions in the two tasks. Participants were faster in accurately completing negative sentences, but they were less accurate when judging novel words to be negative in valence judgment.

The third prediction about word valence inference was supported by the results of *meaning generation*, where the valence of generated meanings reflected the relative emotional valence of the context in which the novel words appeared.

The inconsistent results of *sentence completion* in the immediate post-test and *valence judgment* in the delayed post-test regarding the responses to negative options prompted us to conduct a post hoc analysis to explore whether participants exhibited a preference for options of a specific valence, regardless of accuracy. Rather than calculating the accuracy of responses for each condition and only considering RTs for correct trials, we combined both correct and incorrect trials to look at the proportion of times an option of a particular valence was selected.

Furthermore, while we were mainly interested in the influence of emotional versus neutral context and did not hypothesise for the directionality of the influence of valence, we conducted a post hoc comparison between positive and negative contexts.

Exploratory analysis

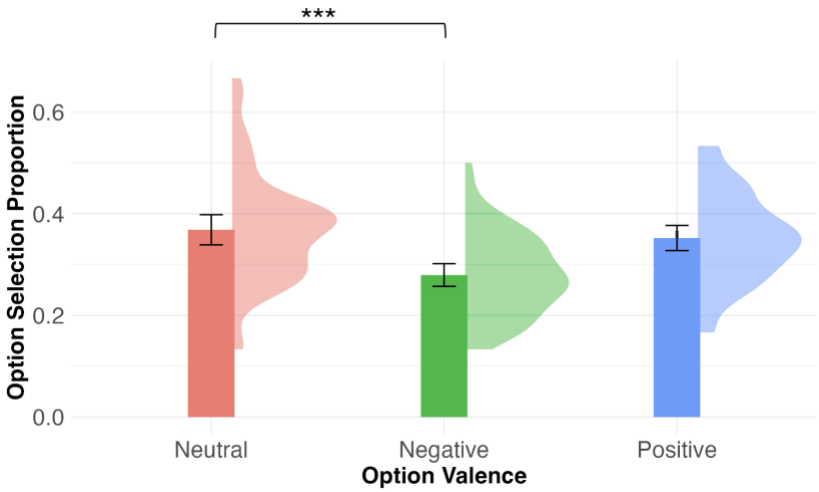
To explore the directionality of the influence of valence, we conducted a post hoc comparison between positive and negative contexts. We found that the difference between the two was significant only in *meaning generation* and *valence judgment*. In *meaning generation*, as compared to novel words appearing in the negative context, participants assigned a more positive meaning to novel words experienced in the positive context ($\beta = 1.08$, $SE = 0.14$, $t = 7.73$, $p < .001$). In *valence judgment*, positive words were judged faster and more accurately than negative words (accuracy: $\beta = 0.33$, $SE = 0.11$, $z = 3.13$, $p = .002$; RT: $\beta = -1.68$, $SE = 0.56$, $t = 3.01$, $p = .003$). The difference in accuracy and RT between positive and negative conditions in *speeded recognition* and *sentence completion* did not reach significance (ps > .05).

We then conducted a post hoc analysis of possible option preference, to better understand the inconsistent results of *sentence completion* in the immediate post-test and *valence judgment* in the delayed post-test regarding the responses to negative options. A linear model was adopted with option

selection proportion as the dependent variable and included the predictor of option valence. In *sentence completion*, option valence was not a significant predictor for the proportion of times an option was selected, $F(2, 225) = 0.07$, $p = .93$. As compared to neutral options (novel words that were encountered in neutral contexts; $M = 0.34$, $SD = 0.08$), there was no difference in preference for the negative options ($M = 0.33$, $SD = 0.07$) or the positive options ($M = 0.33$, $SD = 0.07$; negative vs. neutral: $\beta = -0.0009$, $SE = 0.01$, $t = -0.07$, $p = .94$); positive vs. neutral: $\beta = -0.004$, $SE = 0.01$, $t = -0.36$, $p = .72$). For the analysis of RT, option valence was a significant predictor of log-transformed RT, $\chi^2(2) = 10.22$, $p = .007$. As compared to the novel words previously encountered in neutral context ($M = 4898$ ms, $SD = 2230$), novel words previously experienced in the negative context ($M = 4501$ ms, $SD = 1881$) were selected faster, $\beta = -0.06$, $SE = 0.02$, $t = -3.14$, $p = .002$. Difference between positive ($M = 4838$ ms, $SD = 2483$) and neutral options were not significant, $\beta = -0.02$, $SE = 0.02$, $t = -0.90$, $p = .37$).

In *valence judgment*, option valence was a significant predictor of the proportion of times selected, $F(2, 225) = 20.16$, $p < .001$. Figure 4-5 shows the mean option selection proportion by option valence. As compared to neutral options ($M = 0.37$, $SD = 0.11$), negative options ($M = 0.28$, $SD = 0.08$) were selected less of the time, but the difference between positive ($M = 0.35$, $SD = 0.09$) and neutral options were not significant (negative vs. neutral: $\beta = -0.09$, $SE = 0.01$, $t = -5.96$, $p < .001$; positive vs. neutral: $\beta = -0.02$, $SE = 0.01$, $t = -1.09$, $p = .28$). For the analysis of RT, linear mixed model was used with inverse square root-transformed RT as the dependent variable and included the fixed effect of option valence and random intercepts for participants and stimuli. Option valence was a significant predictor of reaction times, $\chi^2(2) = 8.35$, $p = .02$. As compared to reaction times to the neutral options ($M = 1359$ ms, $SD = 889$), negative options ($M = 1376$ ms, $SD = 722$) were selected significantly slower, but the difference between positive ($M = 1329$ ms, $SD = 775$) and neutral options were not significant (negative vs. neutral: $\beta = 0.87$, $SE = 0.36$, $t = 2.44$, $p = .01$; positive vs. neutral: $\beta = -0.10$, $SE = 0.34$, $t = -0.29$, $p = .77$).

Figure 4-5. Option selection proportion by option valence in valence judgment (exploratory analysis).



Note. The density plots represent the distribution of the mean option selection proportion. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Discussion

Our study examined if and how emotional contexts, varying in valence, influence the learning of novel words, and whether word valence can be inferred from contextual valence. Participants learned novel words embedded in short narrative contexts of either positive, neutral, or negative valence. We found that across the different contexts, participants were able to distinguish target novel words from distractors, both in the immediate and delayed post-tests. As compared to words learned in the neutral context, participants were more accurate in recognising word forms learned in the negative context (in both immediate and delayed post-tests), and the positive context (only in the immediate post-test). Participants were faster in negative sentence completion, but less accurate when making negative valence judgments. We also found that people inferred the valence of novel words from the linguistic context in which they appeared, where the valence of generated meanings reflected the relative emotional valence of the context in which the novel words appeared.

Our findings that participants estimated the valence of novel words after brief exposure aligned with existing research. Following the general correlation between contextual valence and word valence across a large corpus of text (Snefjella & Kuperman, 2016) and the suggestion that the valence of the word might be inferred from its context, Snefjella et al. (2020) and Lana and Kuperman (2023) provided initial experimental evidence for the transfer of valence to novel concrete nouns. Our study supports and extends this finding to adjectives, which tend to be more abstract and emotionally charged. Furthermore, while both Snefjella et al. (2020) and Lana and Kuperman (2023) used a valence rating task, our study employed a novel task that did not probe valence directly. Instead, we asked people to capture the meaning of each novel word in one word. We then obtained the valence of the response words from existing norms. This provided evidence for the implicit transfer of contextual valence to novel adjectives, without having to ask participants to explicitly reflect on valence directly. A different type of interpretation is that participants actively associated the novel word forms with existing words during the reading phase of the study. This might be thought of as akin to second language learning, whereby the difference in valence attributed to a novel word reflects the already known word to which the novel word is associated. While plausible, this does not explain our results entirely. The novel adjectives did not have a predefined meaning that could be simply translated or mapped to a familiar word. Note also that in the meaning generation task, participants were not provided with the narrative contexts. Instead, they had to reflect on the context of each newly learned novel word, and then find an English word that captures this understanding of meaning in context. We suggest, therefore, that the contextual valence experienced during the initial reading lingered and influenced their performance on the post-tests, which is reflected in the meaning generation task. This successful inference of valence from contexts to novel words supports the affective embodiment account, which suggests that emotional content aids in grounding abstract word meanings in affect, allowing for the development of lexical representations from linguistic rather than physical experiences (Vigliocco et al., 2014). This has implications for language acquisition for abstract words, which often rely on emotional cues for meaning since they lack direct sensorimotor connections (e.g., Borghi et al., 2017; Kousta et al., 2011; Ponari et al., 2018).

Our results suggest that both word form and their meanings can be learned from reading short narratives. For word form learning, participants showed above-chance sensitivity to distinguish learned items from distractors in both immediate and delayed *speeded recognition*. This corroborates previous research showing that even a few encounters with a novel word during narrative reading may result in robust learning of visual word forms (e.g., Hulme et al., 2023; Mak et al., 2021). We found that participants were more accurate (but not faster) in recognising words learned in emotional context (both negative and positive) in the immediate post-test. Turning to word meaning, participants showed above-chance accuracy in both *sentence completion* and *valence judgment*. In *sentence completion*, while there was no difference in accuracy, contextual valence was a significant predictor for reaction times for correct trials, and negative items were responded to faster. Together, these results provide further evidence for the affective embodiment account for word meaning acquisition and the role of emotional valence in providing an embodied learning experience in which to anchor abstract meanings (Ponari et al., 2018, 2020; Snefjella et al., 2020). From this view, emotional content of the context aids in grounding abstract word meanings in emotional experiences. Our study focused specifically on adjectives, which tend to be more abstract than nouns and therefore perhaps more dependent on emotional cues that offer a grounding that would otherwise be unavailable from sensorimotor experiences.

While not part of our initial hypotheses, some of our findings point to a negativity advantage. In *speeded recognition*, we found that after a 24-hour delay, words learned in the negative context were recognised more accurately compared to words learned in the neutral context. Similar results were found in *sentence completion*, where participants saw a sentence with a missing word and selected the best completion word option from three novel words that they had encountered in the reading phase (one from each valence condition). We found that as compared to new neutral sentences, participants were faster to complete new negative sentences with the correct word option that was initially encountered in a negative context in the reading phase. The word option that carries a negative connotation might be more attention grabbing and faster to respond to. These negative completion word options might also make other options slower to respond to in non-negative sentence contexts.

This finding suggests a subtle yet significant processing advantage for negatively valenced information (Vaish et al., 2008). Thus, the sustained negativity advantage in delayed *speeded recognition* and the faster processing of negative words in *sentence completion* suggests that negative information is recognised and remembered better (Ortony et al., 1983). The Negative Emotional Valence Enhances Recapitulation model (Bowen et al., 2018) highlights the role of negative valence in enhancing the reactivation of sensory details, consistent with negative events being encoded and retrieved with greater sensory fidelity (Vaish et al., 2008). This contrasts with the positivity advantage suggested by the information density hypothesis (Unkelbach et al., 2008). This might be because a one-day delay in our study and that participants only experienced these novel words in two narratives might not be sufficient for novel words in the positive condition to become integrated into the existing lexicon. In comparison, Snefjella et al. (2020) and Lana and Kuperman (2023) found a consistent positivity advantage for newly learned words, but this was after five exposure opportunities and across a one-week interval. The amount of exposure and its time course might explain differences across studies. Another possible explanation for the deviations from the previous positivity advantage found in Snefjella et al. (2020) and Lana and Kuperman (2023) might be due to the nature of the target words, which are adjectives in the current experiment. From a language evolution perspective, compared to nouns, adjectives are more robustly associated with valence-dependent mutation, suggesting that the meanings of negative adjectives are more differentiated because there are more of them, and they are acquired at a faster rate (Jackson et al., 2023). We might therefore be more adapted to learn novel adjectives that carry a negative connotation as compared to nouns.

Against this, however, findings from our *valence judgment* task appear to contradict the negativity advantage discussed above. Participants were less accurate in categorising novel words as negative when they had previously appeared in a negative context. To reconcile the discrepancy, we propose that the divergence may stem from the explicit nature of the demands introduced by the valence judgment task. While both *sentence completion* and *valence judgment* employed a three alternative forced choice format, with options representing each of the valence conditions, *valence judgment* explicitly asked participants to decide on a valence category (positive, neutral, or negative).

Participants might thus engage in different cognitive processes compared to tasks where valence assessment was not explicit and was instead disguised under the probing of meaning. The difference in task demand was an active decision on our part, as was the ordering of the tasks: as *valence judgment* was the last task administered, the explicit instruction to think about emotional valence could not bias behaviour in earlier parts of the study. For this more explicit assessment of valence knowledge, accuracy was lowest for words previously encountered in negative contexts. Exploratory analysis of the valence category choices made by participants in the *valence judgment* further demonstrated that regardless of response accuracy, the negative category was least selected, and even when selected, responses were slow. This non-negative preference when an explicit valence judgement is required might indirectly support the Pollyanna hypothesis, which describes a universal tendency to use positive language more often (Boucher & Osgood, 1969; Dodds et al., 2015). From this perspective, when participants have to make a valence decision, they are more likely to judge words as positive than negative.

The mixed results in the literature regarding positivity and negativity advantages might partially stem from differences in task demands and experimental paradigms. Tasks vary in complexity and may tap on different processing mechanisms, leading to different observation of results (e.g., Hsiao et al., 2020; Mak et al., 2021). In our study, for example, *speeded recognition* is relatively straightforward, but *valence judgment* performance is tricky to interpret as participants might have different understandings of valence and adopt different benchmarks to judge a word as positive, neutral, or negative. Additionally, the terms “positivity advantage” and “negativity bias” are often loosely employed in the literature and might manifest differently depending on the specific cognitive task. As Unkelbach et al., (2020, p. 119) noted: “It is not *a priori* clear what constitutes an advantage or disadvantage in a given situation.” In lexical decision, for example, positive words are processed faster, suggesting a positivity advantage. However, when semantic decision is concerned, negative stimuli may attract more attention and be processed more slowly, leading to longer processing times which supports an increased vigilance for negative stimuli, indicative of a negativity bias (e.g., Fiske, 1980; Vaish et al., 2008). While there is an emotionality effect, valence effects might be at least partly

task-specific (Crossfield & Damian, 2021). A systematic review exploring these phenomena across various task demands and paradigms would be useful, and future research should aim for precision in defining and contextualising these terms within specific experimental settings.

Overall, our study builds on existing research to investigate the influence of emotional narrative contexts on word learning and extends it to learning adjectives via naturalistic reading. Participants showed evidence of learning after just two exposures, especially for words learned in emotional contexts. Newly learned words also inherited valence from the linguistic context in which they appeared, which holds implications for how emotional knowledge builds during language acquisition. With this in mind, it is worth noting that research on the influence of emotional context in word learning is still in its infancy and mainly targets adults. We know of no study similar to adults that investigates how children learn the emotional properties of new words during independent reading. This represents a knowledge gap, given experience with written language provides opportunities to learn new words and abstract concepts (Nation et al., 2022), and that the valence of existing words can influence children's word learning (e.g., Ponari et al., 2018). While there is evidence for a positivity advantage in younger children, there is also evidence to suggest that the effect dissipates with age (Bahn et al., 2017; Ponari et al., 2018) and that older children write stories which are less positive (Dong, Hsiao, et al., 2024). Hence, it would be interesting to extend our experimental approach to children, which we then present in Chapter 5.

Chapter 5. Emotional Context in Children's Language and Word Learning

Chapter summary:

Analysis of adult language corpus (e.g., Snefjella & Kuperman, 2016) shows that the valence of the word positively correlated with the valence of its surrounding language context, and Chapter 4 has shown that context valence predict word learning in adult experiments. Little is known about whether this extends to children. To address these gaps, we first conducted a computational analysis that quantifies the emotional context of words within a large corpus of children's writing, and the results replicated the pattern of correlations between context valence and word valence seen in adult language. We then conducted a pre-registered word learning experiment to investigate how emotional narrative context shapes the learning of novel adjectives during naturalistic reading. 120 children aged 7 to 11 years from UK primary schools read 15 novel words (such as "*garive"*) embedded in 30 short narratives of either neutral, negative, or positive valence. There were 10 narratives in each valence condition and each novel word appeared twice in two narratives of the same valence. Three immediate post-tests assessed learning. We found that children were able to infer word valence from narrative context, supporting context valence as effective cues for word learning. Children were able to learn novel adjectives from reading short narratives, and older children outperformed younger children. Novel adjectives read in more emotional (positive or negative) contexts were recognised more accurately than those read in neutral narratives. The findings extend previous research conducted using noun concepts and with adults, providing further evidence for affective embodiment in supporting the learning of abstract concepts.

Introduction

In language, emotional valence refers to the pleasantness of a word (Warriner et al., 2013). Large-scale studies have established valence norms by asking adult participants to rate on a scale how they feel while reading a word in isolation. Rated valence is associated with language processing. For

example, in adults, positive and negative words are responded to more quickly in lexical decision tasks than neutral words (e.g., Kousta et al., 2009). Children also show this effect and more generally, valence predicts age-of-acquisition ratings, with positive and negative words being earlier acquired than neutral words (e.g., Ponari et al., 2018). But how do children learn the valence of a new word? One possibility is via the language context it appears in. Emotional context influences word learning in adults (e.g., Dong, Mak, et al., 2024; Lana & Kuperman, 2023; Snefjella et al., 2020) and complementing this evidence, Snefjella and Kuperman (2016) found that the valence of an individual word correlated with the valence of the contexts in which it occurs across a large corpus analysis of adult language. Given these findings, it is plausible that as children experience a new word, they can acquire its valence from the valence of the language contexts in which it occurs. However, no existing work has considered language contexts experienced by children, and relatedly whether the emotional valence of the narrative context serves as a meaningful cue for word learning in children. We aimed to fill these gaps in two ways. We first conducted a corpus analysis to establish whether there is a correlation between word valence and context valence in children's own language and following this, we tested in a word-learning experiment whether context valence is an effective cue as children learn new words.

We focus on school aged children and written language for several reasons. Previous work on the connections between emotion and language has tended to focus on early acquisition (e.g., Nencheva et al., 2023; Ogren & Sandhofer, 2021), yet emotion and language continue to co-develop beyond the preschool years through childhood and adolescence (e.g., Baron-Cohen et al., 2010; Doost et al., 1999). The onset of literacy in mid-childhood (from around 6 years of age) heralds the opportunity to learn via reading, and indeed, reading may provide distinctive learning opportunities, given the decontextualised nature of written language. Readers cannot benefit from cues such as facial expression, gesture, and tone of voice in the way that conversational partners can. Instead, written language itself needs to be crafted so that the situation intended by the writer can be reconstructed in the mind of the reader. As a result, written language tends to be more complex than spoken language, a finding that extends even to books written for young children (Nation et al., 2022). Children's books

are lexically richer than day-to-day conversational language and therefore reading (or being read to) allows children to experience a more diverse range of vocabulary, including less frequent and more complex words (e.g., Dawson et al., 2021; Korochkina et al., 2024). Particularly relevant for our investigation, children's books contain more words that are emotionally arousing (Dawson et al., 2021) and more words that are considered to tap emotions and mental states (Dong & Nation, 2024) than child-directed speech. This means that written language provides a rich substrate from which children have the opportunity to learn emotion vocabulary and word valence. While there is evidence for a general association between reading and social-emotional development (e.g., Batini et al., 2021), whether and how valence influences word learning via independent reading has not been investigated directly, despite reading being the major determiner of vocabulary development, once children can read (Nagy et al., 1985, 1987).

Emotional valence facilitates word learning and word processing

As noted above, emotionally valenced words, whether positive or negative, tend to be acquired earlier than neutral words, especially abstract ones such as *great* and *sad* (Kousta et al., 2011; Ponari et al., 2018). This earlier acquisition may be facilitated by emotional valence acting as a bootstrapping mechanism that aids the learning of abstract concepts (Vigliocco et al., 2014). Bootstrapping refers to the process by which children use existing knowledge and experiences, such as emotional experiences, to help them understand and learn new abstract concepts. Ponari et al. (2020) taught children new abstract words that were either valenced (negatively or positively, e.g., tyranny, crude; karma, insight) or neutral (e.g., trend, proxy), according to the word valence norms by Warriner et al. (2013). Ponari et al. (2020) found that 7- to 9-year-olds showed better learning of the valenced words compared to the neutral words, consistent with the proposal that emotional experience grounds the learning of abstract words and concepts (as described in the affective embodiment account, Vigliocco et al., 2014). These findings align with auditory lexical decision data showing that children respond more accurately to abstract words high (vs. low) on emotional valence (Ponari et al., 2018).

There is some evidence that this emotionality advantage might interact with the age of the children.

Ponari et al. (2018) tested three age groups: children aged 6 to 7, 8 to 9, and 10 to 11 years, with 20 children in each group. The processing advantage of emotionally valenced abstract words was most pronounced in the 8- to 9-year-olds. In contrast, in another auditory lexical decision experiment, Lund et al. (2019) tested children aged 5 to 7 years and found a faciliatory effect of valence only in 6- to 7-year-olds. It is difficult to draw conclusions across these studies due to methodological differences. For example, Lund et al. (2019) included words that were less abstract than those used by Ponari et al. (2018). Other studies have shown diminished effects of valence in words that are less abstract, and for older children (Reggin et al., 2021; Vigliocco et al., 2018). Vigliocco et al. (2018) suggested that the changing emotionality effect with age might indicate a shift from an affective embodied model to a more language-based one for children's abstract word acquisition. On this view, as children's language becomes more sophisticated, familiarity with the distributional properties might play an increasingly important role in complementing affective embodiment to support the learning of abstract words and concepts. Furthermore, linguistic development and social development might also support word acquisition at different developmental stages (e.g., Borghi et al., 2017; Kousta et al., 2011).

However, the role of affective embodiment in abstract word acquisition has been questioned due to inconsistencies and methodological challenges, such as the inconsistent valence effect, and the absence of interaction between valence and concreteness (Bireta et al., 2023). For example, J. M. Kim et al. (2020) found an effect of emotional valence in children but not in adults, with no interaction between valence and concreteness in either group. Additionally, dimensions such as arousal and dominance were not strictly controlled, making interpretation and comparisons across studies difficult (Bireta et al., 2023).

In the presence of a valence effect, the empirical results regarding whether positive or negative valence is more effective in facilitating the acquisition of abstracts words have been mixed (Kauschke et al., 2019). Some studies (e.g., Ponari et al., 2018) reported a positivity effect, which can be explained by the Pollyanna Hypothesis (Boucher & Osgood, 1969), also known as the linguistic

positivity bias where people tend to use more positive language and remember positive things better (Dodds et al., 2015). Yet, Silk et al. (2009) found that children and adolescents show no clear advantage for positive or negative words, and some studies even reported a negativity effect, where children demonstrate better recall for negative words (J. M. Kim et al., 2020). This negativity effect could be due to the greater attention people allocated to negative concepts, which might serve to enhance memory retention (Vaish et al., 2008). This mixed evidence highlights the complexity of emotional word processing in children and how it might also be influenced by factors such as developmental stage, emotional context, and task demands.

Word valence and context valence

In adult language, there is a positive correlation between the valence of individual words and the emotional tone of the contexts in which they appear. For example, Snefjella and Kuperman (2016) analysed the valence, arousal, and concreteness of words and their contexts, using text samples taken from a large corpus of email newsgroup postings (USENET, Shaoul & Westbury, 2013). Ratings of word valence, arousal, and concreteness were taken from norms established by mega studies (Brysbaert et al., 2014; Warriner et al., 2013). A word's context valence was defined as the aggregate valence of the five content words immediately before and after the word, calculated across the corpus. Snefjella and Kuperman (2016) found medium-to-strong correlations between the valence, arousal, and concreteness of individual words, and the valence, arousal and concreteness of the contexts in which they occurred. To further corroborate this finding, Kuhlmann et al. (2017) used corpus-derived co-occurrence statistics as a measure of association strength and asked whether a neutral word would be judged more as positive or negative based on its association strength with valenced words. Using a valence decision task, they found that if a neutral word more often appears together with positive words in natural language, participants were faster in judging this word as positive than negative. This suggests that valence may function as a "semantic super-feature" that is represented in the patterns of association in semantic networks. For adults, at least, this implies that a word's emotional valence can be inferred from its contextual experience.

These findings may have implications for children's language development. As children grow older, increased exposure to written language through independent reading offers abundant opportunities to encounter new words. In natural reading, texts do not usually provide explicit definitions of new words. Instead, the meaning (and valence) of new words needs to be inferred from context. While word- and context-level affective statistics correlate in adult language, the extent of these correlations have not been established in children's language. There might be differences. For example, Wild et al. (2013) noted that the linguistic context in children's books often focuses on more tangible aspects such as physical descriptions and object locations, as distinct from the abstract and temporal elements that are prevalent in adult language. Our first aim in this paper is to establish the correlation between word valence and context valence across a large sample of children's language, using the methodology of Snefjella and Kuperman (2016).

To preface the results of Study 1, we analysed children's written language and found a positive correlation between context and word valence. This finding sets the scene for Study 2 which took an experimental approach to test the causal hypothesis that children are sensitive to context valence such that this serves as an effective cue for word learning, building on previous studies with adults. There is consistent evidence to indicate that adults can infer valence for new words based on the context in which they are embedded (e.g., Dong, Mak, et al., 2024; Driver, 2021; Frances et al., 2020; Lana & Kuperman, 2023; Snefjella et al., 2020). For example, Snefjella et al. (2020) asked native speakers of English to learn nine novel nouns (e.g., *plurk* as a substitute for "the real word"), each embedded in five short passages that were designed to be neutral, negative, or positive. In the subsequent post-tests, participants showed clear evidence of word learning, with the positive condition resulting in the best learning outcomes. Importantly, all the novel words acquired emotional connotations corresponding to the contexts in which they appeared. Dong, Mak, et al. (2024) adopted a similar paradigm and extended to the learning of adjectives and again found that words experienced in emotionally valenced contexts were better learned than those experienced in neutral contexts, and that participants were able to infer the valence of the novel words based on its context. While a general effect of an emotionality advantage holds across studies, note that there is some inconsistency as to whether there

is a difference between positive and negative conditions. Differences across studies might reflect differences in experimental design, and whether the learning is incidental or via explicit teaching.

A study by Ponari et al. (2020), introduced earlier, investigated similar themes in children. They taught 7- to 10-year-olds new abstract words that were neutral, negative or positive. One group of children were taught the words in a condition that emphasised emotional information, while the other group were taught the new words via more encyclopaedic, non-emotional information. Abstract words with emotional valence were better learned, consistent with the idea that emotional valence serves as a bootstrapping mechanism for acquiring abstract concepts. While they found no overall difference in learning for the two types of teaching approach, context was operationalised very differently to the method used in the adult word learning studies reviewed above (Dong, Mak, et al., 2024; Driver, 2021; Frances et al., 2020; Lana & Kuperman, 2023; Snefjella et al., 2020), and even their 'non-emotional' condition included many valenced words, thereby creating a valenced context from which children could potentially learn, regardless of the emotion vs non-emotion encyclopaedic teaching strategy. In addition, Ponari et al. (2020) focused on explicit and direct teaching, repeated in multiple sessions extending across a week. This is quite different to the incidental learning targeted in the adult word learning experiments, and more importantly, it is quite different to how children experience new words in narrative contexts when reading independently.

The present research

In summary, the emotional context provided by text is associated with word valence and word learning in adults, but we do not know whether this extends to children. It is important to address this knowledge gap given that written language is a primary avenue from which children learn new words and abstract concepts, especially for words that are not likely to be encountered in day-to-day conversations. There is also a need for experiments to more closely resemble naturalistic reading, where words are experienced and presumably learned via reading itself, rather than via direct instruction and explicit definitions. To address these issues, we conducted two studies: (1) a corpus analysis investigating the emotional properties of words and their contexts in children's language and

(2) a word-learning experiment testing whether children learn the emotional properties of new words when reading narrative.

Study 1: Word Valence and Context Valence in Children's Narrative Writing

In adult language, there is a positive correlation between the rated valence of an individual word and the valence of its context, defined as the aggregate emotional tone of the surrounding contexts it occurs in, and more generally, studies utilising large language corpora have confirmed that words occurring in similar emotional contexts tend to share semantic properties (Kuhlmann et al., 2017; Snefjella & Kuperman, 2016). These prior studies used email newsgroup corpora, comprising postings written by adults (USENET, Shaoul & Westbury, 2013). To extend these findings to children's language, Study 1 analysed a large corpus of stories written by children. We focused on valence and context valence, but we also calculated and reported two other variables (arousal and concreteness) included in the study by Snefjella and Kuperman (2016), for completeness and future reference.

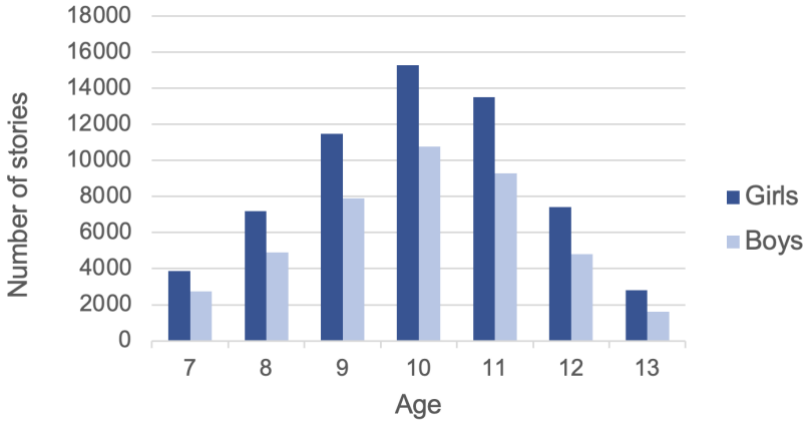
Method

Materials

We used a large corpus of stories written by 7- to 13-year-old children and submitted to a UK children's writing competition (BBC 500 Words) in 2019. BBC 500 Words is an annual national competition for children. This has been running for over 10 years with the same format in which children (aged 5-13 years until 2020, and then aged 5-11 when the competition returned in 2023) across the UK were invited to submit a story on any theme or topic, so long as the word count was not greater than 500 words. This resource therefore provides a naturally occurring language sample not generated for assessments or in response to prompts or cues, and this allows us to analyse children's own language, free from constraints on time, instruction, or topic. For this study, we selected all stories submitted in 2019 (N = 107,273; approximately 55-million word tokens). We excluded stories written by children aged 6 and under as the number of entries was relatively low; it also meant that

the language sample came from children similar in age to those who participated in Study 2. The analysed corpus comprised 103,541 stories, 46,697,930 word tokens and 218,914 word types in total ($M_{words/story}$ = 451.01, $SD_{words/story}$ = 92.16). Figure 5-1 shows the distribution of stories by age and gender. More girls (59.44%) contributed stories than boys (40.56%), and most stories came from 9- to 11-year-olds.

Figure 5-1. Number of stories analysed by age and gender.



Following the methodology used by Snefjella and Kuperman (2016), the semantic properties of words and their contexts in the corpus were taken from existing mega study norms. Valence and arousal ratings for English lemmas (N = 13,915) were taken from Warriner et al. (2013) and concreteness ratings for words (N = 40,000) from Brysbaert et al. (2014). We enhanced the set of affective norms by assigning the value of valence and arousal given to the lemma (e.g., sing) to all its inflected word forms (i.e., sang, sung, singing), based on the word form-lemma matching data in the Child and Young People's Lexicon (CYP-LEX; Korochkina et al., 2024). This increased the number of word forms with associated valence scores to 30,872.

Procedure

Unless otherwise specified, our analysis procedure closely followed the procedure outlined in Snefjella and Kuperman (2016). A word's context was defined as the five content words that immediately preceded it, and the five content words that immediately followed it in each story. The target word itself was not considered as part of its context. To compute the context for each word, we first converted all characters in children's stories to lowercase and removed function words and punctuation. For each occurrence of a word, we calculated how positive, arousing, or concrete the word is, where words within each context (excluding the target word) were matched with their ratings from the expanded norms based on Warriner et al. (2013) and Brysbaert et al. (2014). We included only those contexts where ratings were available for at least three words. We then averaged the valence, arousal, and concreteness of contexts across all occurrences of each word in the corpus (N = 25,348 words). We further excluded 965 words (4%) whose overall context valence, context arousal, or context concreteness were more than 3 SDs above or below the mean of the respective variable, leaving 24,383 words for which we had semantic estimates for both individual words and their contexts.

The three contextual variables (valence, arousal, and concreteness) served as indices of the overall tendency of a word to occur in positive, exciting, or concrete contexts. We then calculated Pearson correlation coefficients between the values for individual words and for the contexts in which they occurred.

Results

Summary statistics for the words and their associated contexts are presented in Table 5-1. The context variables are normally distributed, as seen in Figure 5-2.

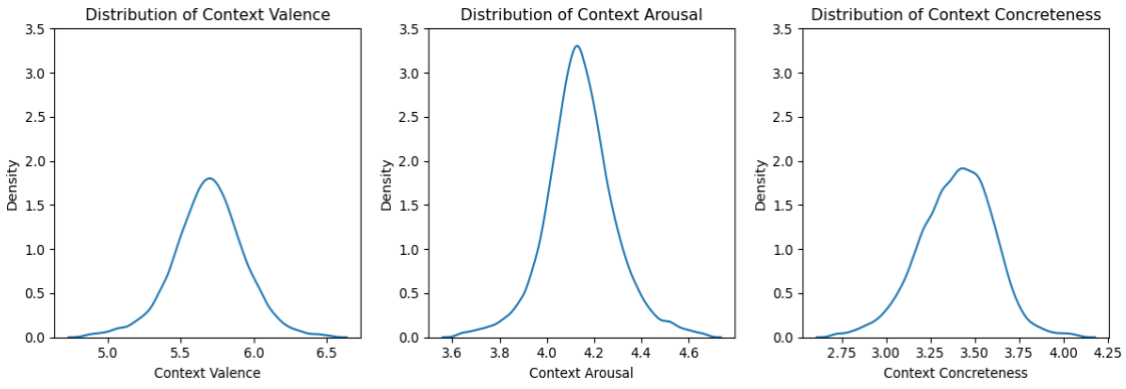
Table 5-1. Summary statistics of word-level and context-level properties in the children's writing corpus (N = 24,383 words).

	Mean	SD	Min	Max	Skew	Kurtosis
Word Valence	5.13	1.27	1.40	8.53	-0.35	-0.28
Word Arousal	4.20	0.90	1.60	7.79	0.49	0.05
Word Concreteness	3.46	1.01	1.12	5.00	-0.15	-1.21
Context Valence	5.69	0.25	4.83	6.54	-0.12	0.62
Context Arousal	4.14	0.15	3.62	4.68	0.11	1.08
Context Concreteness	3.40	0.21	2.69	4.09	-0.19	0.27

Note. Valence and arousal were rated from 1-9, with 1 being very negative and 9 being very positive; Concreteness was rated from 1-5, with 1 being very abstract and 5 being very concrete.

Figure 5-2. Density plots showing the distribution of (a) Context Valence, (b) Context Arousal, and (c) Context Concreteness.

(a) (b) (c)



Note. Valence and arousal were rated from 1-9, with 1 being very negative and 9 being very positive; Concreteness was rated from 1-5, with 1 being very abstract and 5 being very concrete.

The correlations between context-level and word-level variables are shown in Table 5-2. We found that the values for individual words were moderately correlated with the contexts in which they occurred across all three properties (valence $r = 0.48$; arousal $r = 0.35$; concreteness $r = 0.50$)[2].

Table 5-2. Correlation matrix of word-level and context-level properties in children's language.

	Word Valence	Word Arousal	Word Concreteness	Context Valence	Context Arousal	Context Concreteness
Word Valence	-					
Word Arousal	-0.20***	-				
Word Concreteness	0.08***	-0.16***	-			
Context Valence	0.48***	-0.12***	0.01*	-		
Context Arousal	-0.13***	0.35***	-0.12***	-0.26***	-	
Context Concreteness	0.04***	-0.12***	0.50***	-0.06***	-0.17***	-

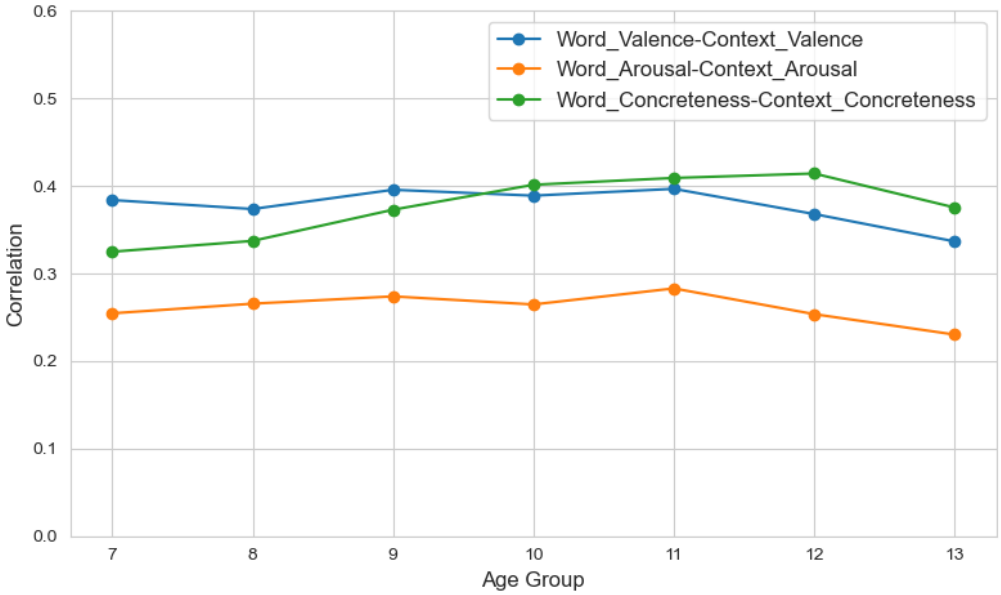
Note. $p < .001$ ***, $p < .01$ **, $p < .05$ * based on Pearson's correlation test.

We also explored if the correlations between word valence and context valence change as a function of children's age, plotted in Figure 5-3. The x-axis represents the different age groups, from 7 to 13 years old. The y-axis is the correlation coefficients between the three pairs of word- and context-level semantic variables (represented in three colours) at each age groups. From observation of the lines,

[2] For comparison, the correlations for adult language reported in Snefjella and Kuperman (2016) are: Word_Valence-Context_Valence: 0.58; Word_Arousal-Context_Arousal: 0.48; Word_Concreteness-Context_Concreteness: 0.72

the pattern of correlations is similar across all age groups, although the correlation at each age group is a little lower[3] than overall correlation coefficients reported in Table 5-3.

Figure 5-3. Correlation between word- and context-level semantic variables across age groups.



Exploratory analysis of adjectives use

To explore the use of adjectives in children's writing, we conducted an exploratory analysis to examine their occurrence across different age groups and genders. We used the spaCy package in Python for automatic part-of-speech tagging. The tokenised corpus was processed to identify adjectives within the text, and for each story, we calculated the percentage of adjectives relative to all content words, excluding stop words and punctuations. The data were then grouped by age and gender, and the mean percentage of adjectives was calculated for each group.

[3] This is perhaps due to the smaller subgroup sample sizes.

The analysis revealed that adjectives accounted for an average of 7.45% of content words in children's writing. The use of adjectives remained relatively stable across the age groups, ranging from 7.32% to 7.81%. In stories written by girls, 7.56% of content words were adjectives, compared to 7.27% in stories written by boys.

Discussion

Study 1 asked whether the patterns of correlation between word- and context-level valence, arousal, and concreteness observed in adult corpora (Snefjella & Kuperman, 2016) is also seen in children's language. Our findings are clear in showing a similar pattern of correlations in a corpus of children's writing. Both word valence (5.13) and context valence (5.69) were slightly above the mean of 5, which shows greater usage of more positive words and contexts. Not only do people have more unique words for positive phenomena and events (Warriner & Kuperman, 2015), but also that the way events are discussed by adults (Snefjella & Kuperman, 2016) or written about in children's stories tends to be more optimistic than pessimistic.

Our correlation coefficients are slightly smaller than those observed in adults, which could be partly due to the different types of corpora used. The non-fictional email correspondences for adults were more likely to focus on factual information and adhere to conventional patterns of communication. These texts might feature more consistent and predictable vocabulary and sentence structures, reflecting the denser and more mature semantic networks of adults. This likely led to stronger correlations. In contrast, the fictional stories written by children might contain more imaginative and less predictable language, with creative and varied expressions and unusual word combinations that were less tied to conventional language use. This variability, coupled with the ongoing development of children's language skills, might account for the weaker correlations between word-level and context-level semantic variables observed in their data.

Moreover, observations from the graphs indicate that the pattern of correlations between word and contextual properties was similar across the age range (7-13 years). This suggests that the word-

context relationship is stable within this developmental period, potentially providing reliable cues that help children infer the valence of novel words and support their word learning. Additionally, the exploratory analysis revealed that adjectives accounted for 7.45% of all content words, with this percentage remaining relatively stable across the age range. Adjectives play a crucial role in enriching vocabulary and developing conceptual understanding, allowing children to describe, measure, classify, and compare objects and events - skills foundational to learning across various subjects (Davies et al., 2022; Griffin et al., 2004). Understanding how children use and acquire adjectives offers valuable insights into broader language development processes. This focus is particularly relevant to Study 2, which investigates how children learn novel adjectives from valenced context.

Study 2: Children's Learning of Novel Adjectives from Emotional Narrative Context

Having established a correlation between word valence and context valence in children's language in Study 1, we then asked whether children infer the valence of novel words from narrative context, as they encounter novel words while reading. We used 7- to 11-year-olds as these children are past the initial stages of reading acquisition and are reading independently; it is also the age range that coincides with an increase in emotion vocabulary (Baron-Cohen et al., 2010; Dong & Nation, 2024) and sensitivity to word valence in lexical decision (Ponari et al., 2018). Previous studies on adults investigated the learning of novel words from emotional context. For example, Snefjella et al. (2020) and Lana and Kuperman (2023) investigated the learning of novel nouns, while Dong, Mak, et al. (2024) focused on adjectives. Lana and Kuperman (2023) found that the inference of valence based on context was only evident for concrete, but not abstract nouns. Given that adjectives tend to be more abstract and their meanings more context-dependent (Davies et al., 2020; Dawson et al., 2021), we chose to focus on adjectives, for which the effect of context valence might be more important. Moreover, the exploratory analysis of study 1 shows that adjectives accounted for 7.45% of all content words, making them an important grammatical class to investigate. Following the paradigm developed by Dong, Mak, et al. (2024), we focused on learning novel adjectives and used a naturalistic reading procedure to investigate the effect of context valence on word learning in children.

We compared word learning after reading novel adjectives embedded in narratives that provided neutral, negative, or positive contexts. Based on findings with adults (Dong, Mak, et al., 2024), and the results of Study 1, we aimed to test the hypothesis that context valence facilitates children's novel word learning, and our predictions were as follows:

1. Children would learn novel words from reading short narratives,
    a. and older children will learn better than younger children.
2. Words experienced in more emotional (positive and negative) contexts would be learned better than words experienced in neutral contexts.
3. Children would infer a valence for the novel words from the linguistic context in which the words appeared.

Method

Design

The experiment was based on a scenario of an alien learning and practising English who occasionally included some 'alien' words in its diary. These served as the novel words in the experiment. In the reading phase, children read 15 novel words embedded in 30 short narratives. The short narratives were of either neutral, negative, or positive valence. In the test phase, children completed several tasks to assess their learning of the novel words.

There was one categorical independent variable, context valence, with three levels: neutral, negative, and positive. This was manipulated within-participant. Age (treated as a continuous variable) served as a covariate. Responses to each post-test served as the outcome measure. The study, including the sample size, hypotheses, exclusion criteria, and analysis plan, was pre-registered ahead of data collection (https://doi.org/10.17605/OSF.IO/MVZS4). Data and analysis scripts are available on OSF (http://tiny.cc/rainy_thesis_chapter4and5).

Participants

In total, 120 children (59 girls, 61 boys) were recruited from Year 3-6 classrooms in primary schools in the UK. Their ages ranged from 7 to 11 years old (M_{age} = 9.70, SD_{age} = 1.12). All children were enrolled in primary schools with English as the language of instruction. Fifty-five children (45.8%) had at least one family member who spoke to them in a language additional to English. Informed opt-in or opt-out consent (based on the preference of each participating school) was obtained from parents or guardians, and children provided verbal assent before starting the experiment. Following our pre-registered exclusion criterion, two children were excluded on the basis of poor reading (defined as 1.5 SD or more below the mean of the standardisation sample on the Test of Word Reading Efficiency Second Edition (TOWRE–2; Torgesen et al., 2012). The final sample comprised 118 children (57 girls, 62 boys; M_{age} = 9.71, SD_{age} = 1.12). There were 30 children from Year 3, 34 from Year 4, 27 from Year 5, and 27 children from Year 6.

Materials

We created 30 short narratives ($M_{word\ count}$ = 18.00, $SD_{word\ count}$ = 2.58) of either neutral, negative, or positive valence (10 narratives in each condition). These were adapted from a pool of 70 narratives devised by Dong, Mak, et al. (2024). The sentiment of each narrative was estimated using a *Bidirectional Encoder Representations from Transformers (BERT)* model, a transformer-based machine learning technique for natural language processing (Devlin et al., 2019). *BERT* considers the entire context of words in a sentence, rather than one word at a time, allowing it to capture emotional nuances in the narratives. The pool of 70 narratives were also rated by 19 native English adult speakers. Both approaches showed that context valence differed significantly across the three valence conditions, with narratives in the positive condition showing the most positive valence, followed by neutral, and then by negative. Narratives were matched across conditions for mean length of utterance. To ensure that the 30 narratives were age-appropriate, they were checked by four native English speakers (including one parent and one speech language therapist) with extensive experience working with children. A sample set of narratives is shown in Table 5-3 (see Appendix 5 for all

narratives).

There were 30 novel words, 15 of which served as target novel words and the others as distractors in the post-tests. Each of the 15 target novel words was embedded twice in two narratives in the same valence condition (see Table 5-3). The novel words were all 6 letters long and did not have a base meaning. They were created to have a nonword stem plus an adjective suffix, for example, the nonword stem *roif-* and the adjective suffix *-ic* led to the novel word *roific*. We chose five adjective suffixes from a list of suffixes with high diagnosticity values for adjectives, as calculated by Ulicheva et al. (2020). The novel words had no replacement orthographic neighbours, according to NWatch (Davis, 2005). Assignment of novel words to the context valence condition was counterbalanced such that a novel word appearing in the positive context for one child appeared in the neutral or negative context for other children. The post-test distractors were the same across children. See Appendix 5 for all novel words used.

Table 5-3. An example narrative for each valence condition.

Context Valence	Examples
Neutral	I went to the roific library today. To get there, I turned left at the roific house.
Negative	I had a lenful argument with my classmate today. Their lenful words hurt my feelings.
Positive	I won a merous prize in writing today. I love the feeling of writing merous stories.

Procedure

Consent forms were distributed by schools, and children who had parental consent (or absence of opt-out consent) were invited to take part. On the day of data collection, the experimenter first explained the purpose of the research and gave a short presentation of what participation would involve to the whole class. The experiment was conducted in a one-to-one setting in a quiet room, adjacent to the

main classroom. The experiment was programmed and hosted on Gorilla (www.gorilla.sc, Anwyl-Irvin et al., 2020). It was completed in a single session taking approximately 30 minutes. Stickers related to the theme of the study were used to keep children's attention and engagement.

After providing verbal assent, the child was introduced to the background of the study, which was structured around the premise of an alien attempting to learn English who, when writing their diary, occasionally replaced English words with words from its own language. Before they read the alien's diary, they completed the sight word reading efficiency component of the TOWRE in which they read as many words from a card as they could in 45 seconds, according to the manual instructions.

Next, they read the alien's diary on the experimenter's laptop via Gorilla. They saw novel words in neutral, negative, or positive narratives, presented in a random order. Each child experienced 15 novel words embedded in 30 short narratives, evenly split into two blocks of 15. They were told to pay attention to these novel words, but no further instruction was provided.

Immediately after reading the narratives, word learning was assessed via three outcome measures, administered in a fixed order: *speeded recognition*, *valence judgment*, and *meaning generation*. Several practice items were provided before each task began.

For *speeded recognition*, modelled after lexical decision, children identified whether they had previously seen a presented letter string. Each trial began with a fixation cross displayed for 250 ms, followed by the letter strings, during which the children pressed buttons on the keyboard to make a judgment (the keys M and Z were marked with a tick and a cross respectively). Accuracy of the response was recorded. Each child responded to 30 items in total (15 target novel words and 15 distractors), presented in a random order in a single block.

They then completed a *valence judgment* task in which they made a categorical judgment to indicate whether the meaning of a novel word (experienced in the reading phase) was neutral, negative, or positive. Each word was first displayed for 300 ms before the three options appeared. Children made a judgment by clicking on what they thought was the most suitable option. The next trial started as soon

as a response was recorded. Accuracy was scored and there were 15 trials in total, presented in a random order in a single block.

The final task was *meaning generation*. In each trial, children were shown one of the 15 newly learnt words (presented in a random order) along with the diary narrative context in which they appeared. They were invited to say one English word that they considered to correspond with the meaning of the novel word. This was typed into interface by the experimenter. The words produced were cross-referenced with norms of valence for English lemmas (Warriner et al., 2013) and these values were used to assign a valence score to each response.

At the end of the session, children completed the phonemic decoding subtest of TOWRE, and they reported their basic demographic information, including age and gender.

Statistical analyses

As set out in the pre-registration, we fitted mixed-effect models with random effects of participants and stimuli for each dependent variable, using the *lme4* package (Bates et al., 2015) in R (R Core Team, 2022). For all full models DV ~ ContextValence + Age + (1+ ContextValence | participant) + (1+ ContextValence | item) that failed to converge, we followed Mak, Curtis, et al. (2023) by simplifying the random-effect structure using the R package *buildmer* (Voeten, 2023). A likelihood ratio test was used to compare the full model to the reduced model to assess whether including the fixed factor ContextValence significantly improved model fit.

For the dependent variable sensitivity *d'* in speeded recognition, a one-sample t-test compared the sample mean to the population mean 0, to test for the overall learning effect. For the dependent variable accuracy for valence judgment, a one-sample t-test compared the sample mean to the chance level, which is 0.33, to test for the overall learning effect.

The fixed effect was context valence (neutral, negative, positive). This was dummy-coded, with the neutral condition serving as the reference level. This yielded two comparisons: neutral vs. positive

and neutral vs. negative. Models were fitted using maximum likelihood estimates. Following these confirmatory analyses that compared neutral versus emotional contexts (both positive and negative), we also explored whether there was a difference between positive and negative contexts, and how this might interact with children's age and gender.

Results

Descriptive statistics are provided in Table 5-4. We first present the confirmatory analyses based on the hypotheses, and then the exploratory analyses.

Table 5-4. Descriptive statistics (mean and standard deviation) for each outcome measure.

	Speeded Recognition	Valence Judgment	Meaning Generation
Variable	Accuracy	Accuracy	Valence score
Neutral	0.67 (0.47)	0.36 (0.20)	5.90 (0.82)
Negative	0.75 (0.43)	0.42 (0.22)	3.27 (0.64)
Positive	0.73 (0.44)	0.34 (0.19)	7.31 (0.50)

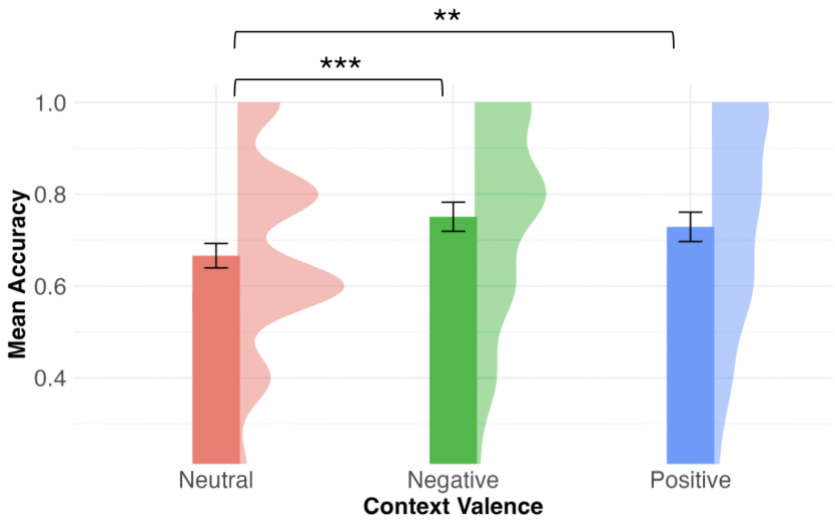
Note. Accuracy is a probability from 0 to 1; Valence score is a continuous value between 1-9, with 9 being very positive, and 1 being very negative.

Confirmatory analyses

Novel word learning was assessed via *speeded recognition*. We first computed sensitivity from raw responses to the speeded recognition task (number of hits, false alarms, correct rejection, and misses) using the *dprime* function in psycho package (Makowski, 2018). A one-sample t-test provided clear evidence that children could distinguish learned items from distractors, $M_{d'} = 1.04$, 95% CI = [0.91, 1.17], $t(117) = 15.87$, $p < .001$. We then compared novel words learned across the different valence

conditions. Figure 5-4 shows the mean recognition accuracy by context valence. A binary logistic mixed effect model was built with context valence and age as the predictors and included the random intercepts for participant and item. From the likelihood ratio test, context valence was a significant predictor for recognition accuracy, $\chi^2(2) = 12.40$, $p = .002$. As compared to words in the neutral context ($M = 0.67$, $SD = 0.47$), children were more accurate in recognising words experienced in the negative context ($M = 0.75$, $SD = 0.43$) and the positive context ($M = 0.73$, $SD = 0.44$; negative vs. neutral: $\beta = 0.47$, $SE = 0.14$, $z = 3.41$, $p < .001$; positive vs. neutral: $\beta = 0.33$, $SE = 0.13$, $z = 2.46$, $p = .01$). Age was also a significant predictor for recognition accuracy, $\chi^2(1) = 26.38$, $p < .001$, with older children performing better on the task than younger children, $\beta = 0.40$, $SE = 0.07$, $z = 5.36$, $p < .001$.

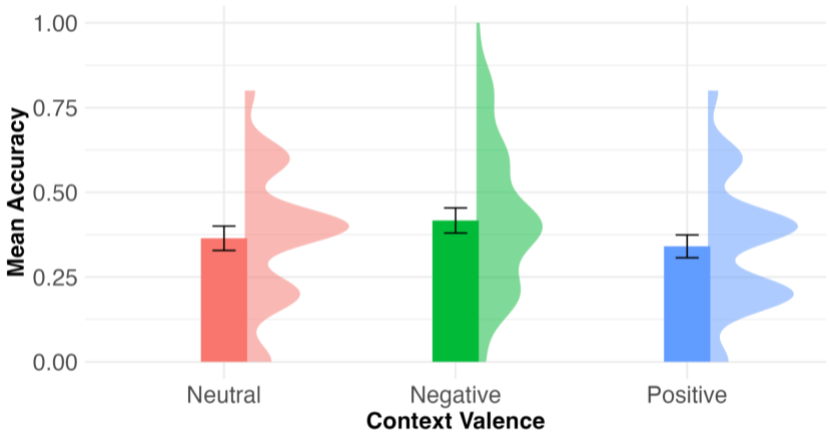
Figure 5-4. Mean recognition accuracy (proportion correct) of novel words experienced in each context valence condition.



Note. The density plots represent the distribution of the mean accuracy in each condition. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Novel word valence learning was first assessed via *valence judgment*. A one-sample t-test provided evidence that children performed above chance, $M_{accuracy}$ = 0.37, 95% CI = [0.35, 0.39], $t(117) = 4.48$, $p < .001$. We then compared novel words learned across the different valence conditions. Figure 5-5 shows the mean valence judgment accuracy by context valence. A binary logistic linear model was built with context valence and age as the predictors. Context valence was a significant predictor of valence judgment accuracy, $\chi^2(2) = 7.67$, $p = .02$. As compared to novel words learned in the neutral contexts ($M = 0.36$, $SD = 0.20$), there is no significant difference in the performance for novel words learned in the negative condition ($M = 0.42$, $SD = 0.22$), or the positive condition ($M = 0.34$, $SD = 0.29$; negative vs. neutral: $\beta = 0.22$, $SE = 0.12$, $z = 1.85$, $p = .06$; positive vs. neutral: $\beta = -0.10$, $SE = 0.12$, $z = -0.85$, $p = .39$. Age was a significant predictor for valence judgment accuracy, $\chi^2(1) = 4.03$, $p = .02$, with older children performing better on the task than younger children, $\beta = 0.09$, $SE = 0.04$, $z = 2.01$, $p = .04$.

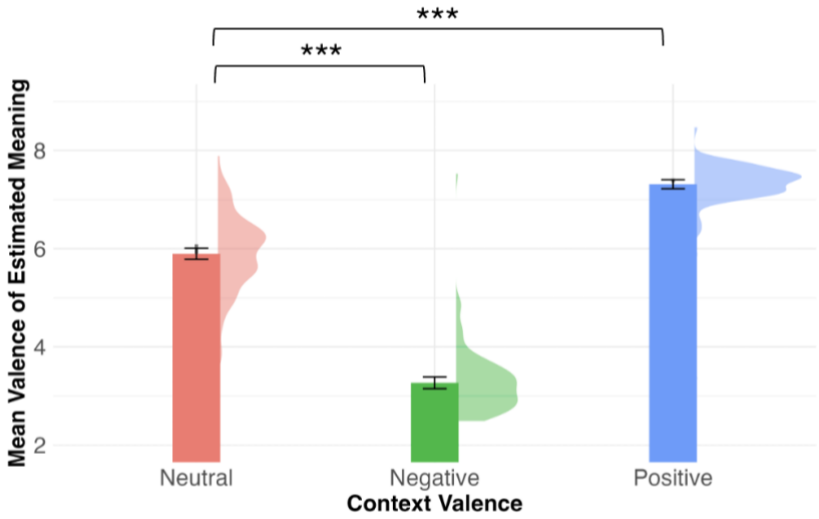
Figure 5-5. Mean valence judgment accuracy (proportion correct) of novel words experienced in each context valence condition.



Note. The density plots represent the distribution of the mean accuracy in each condition. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Novel word valence learning was also assessed via *meaning generation*. 1642 out of 1770 (93%) responses had associated valence scores listed in Warriner et al.'s (2013) norms. Some examples of generated words included *nice, happy, smart, scary, bad, terrible, clean, large*. Majority of generated words are adjectives. A complete list of participant's production can be accessed through the project's OSF page. Figure 5-6 shows the mean valence score of generated meaning per participant by context valence. We built a linear mixed model with the valence score as the dependent variable and included the fixed effect of context valence and age, and random intercepts for participant and item. The likelihood ratio test shows that context valence was a significant predictor of estimated valence scores, $\chi^2(2) = 153.22$, $p < .001$. Compared to novel words appearing in the neutral context ($M = 5.90$, $SD = 0.82$), children assigned more negative meanings to novel words experienced in the negative context ($M = 3.27$, $SD = 0.64$), and more positive meanings to novel words experienced in the positive context ($M = 7.31$, $SD = 0.50$; negative vs. neutral: $\beta = -2.69$, $SE = 0.12$, $t = -22.65$, $p < .001$; positive vs. neutral: $\beta = 1.37$, $SE = 0.12$, $t = 11.59$, $p = .003$). Age was also a significant predictor for estimated meaning valence, $\chi^2(1) = 7.06$, $p = .008$. Older children generated word meanings that were less positive than younger children, $\beta = -0.08$, $SE = 0.03$, $z = -2.68$, $p = .009$.

Figure 5-6. Mean valence scores (as obtained from existing norms) of generated meanings of novel words experienced in each context valence.



Note. The density plots represent the distribution of the mean valence of the meaning generated. Error bars represent 95% within-participant confidence intervals. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. Greater valence represents greater positivity.

To sum up the results of the confirmatory analyses, our hypothesis that emotional valence facilitates children's novel word learning was largely supported. The first prediction that children can learn novel words from reading short narratives was supported by the above chance performance in *speeded recognition*. Age was a significant predictor in all models such that older children performed better than younger children in *speeded recognition* and *valence judgment*. They also generated word meanings that were less positive than younger children in *meaning generation*. Furthermore, as evidenced in speeded recognition, children showed better learning in emotional (positive and negative) contexts, supporting the second prediction, although we found no such evidence in valence judgement. Turning to the third prediction that children would infer a valence for the new words from the linguistic context in which they appeared. This was supported by the above chance performance in *valence judgment* and in *meaning generation*, where the valence of generated meanings reflected the respective emotional valence of the context in which the novel words appeared.

Exploratory analysis

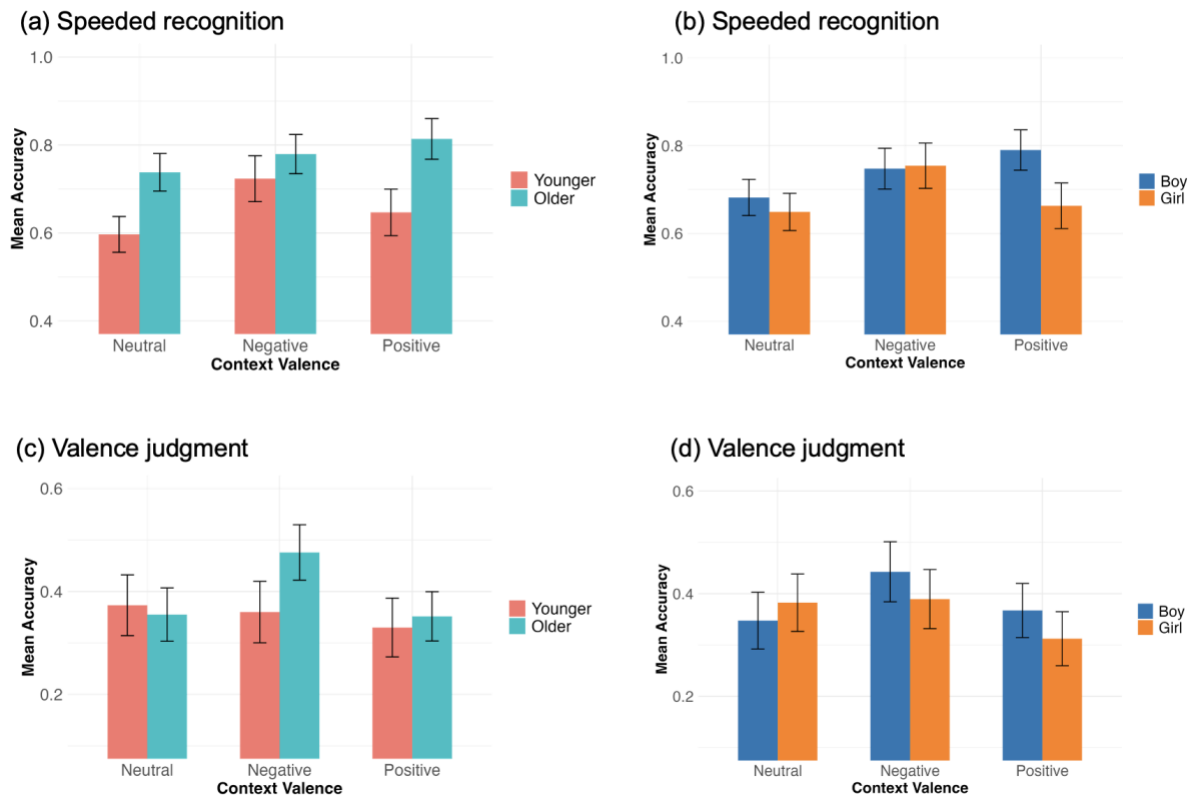
While we were mainly interested in the influence of emotional versus neutral context, we conducted a post hoc comparison between positive and negative contexts to complement the confirmatory analyses. This would allow us to explore whether the emotionality advantage observed might be driven by positivity advantage or a negativity advantage or both. There was no significant difference between positive and negative contexts in *speeded recognition*, $p > .05$ (Figure 5-4). In *valence judgment*, however, words read in the negative contexts were judged more accurately compared to words read in the positive contexts, $\beta = -0.33$, $SE = 0.12$, $z = -2.70$, $p = .007$ (Figure 5-5). In *meaning generation*, compared to novel words encountered in the negative narratives, children assigned a more

positive meaning to novel words experienced in the positive condition, β = 4.06, *SE* = 0.12, *t* = 41.76, *p* < .001(see Figure 5-6).

Moreover, we further analysed whether demographic factors such as age and gender of the children influenced their adjective learning in emotional contexts. For age, the confirmatory analyses showed that older children performed better than younger children, we explored whether this held across the different emotional contexts. To do this, we performed a median split based on age (Median Age = 9.67) with 60 children in the younger age group (*M* = 8.79, *SD* = 0.61) and 58 children in the older age group (*M* = 10.67, *SD* = 0.60). For this exploratory analysis, we chose to present descriptive interpretations based on the visual inspection of the graphs (e.g., whether the confidence intervals overlap) rather than conducting statistical analyses, mindful that our exploratory analyses might be underpowered to test for an interaction, and that some models might not converge when the interaction term was included. Speculatively therefore, Figure 5-7(a) shows that older children recognised novel words read in neutral and positive contexts better than younger children, but there were no age differences for words read in negative contexts. Figure 5-7(c) shows that older children were more accurate in judging a word read in the negative context as negative, but there were no age differences for words read in positive or neutral contexts.

Finally, we explored whether children's gender might influence their performance on various tasks. There were 57 girls and 61 boys. As above, we chose to make descriptive observations based on the graphs. In *speeded recognition,* boys recognised more words in the positive context than girls (Figure 5-7(b)). In *valence judgment*, the confidence intervals overlapped indicating no evidence for gender differences (Figure 5-7(d)).

Figure 5-7. Mean *speeded recognition* accuracy (a-b) and mean *valence judgment* accuracy (c-d) as a function of Context Valence and Age Group or Gender.



Note. Error bars represent 95% within-participant confidence intervals.

Discussion

This pre-registered word learning experiment investigated the influence of emotional narrative context on children's learning of novel adjectives. We found that children can learn novel words from reading short narratives, as seen in the above chance performance in *speeded recognition*. Words read in more emotional (positive or negative) contexts were recognised more accurately than those read in neutral narratives. Older children learned better than younger children. Children were also able to infer word valence from context, as seen in the above chance performance in *valence judgment,* as well as in *meaning generation* where the valence of generated meanings reflected the respective emotional valence of the associated context.

That children can learn novel valenced words through reading short narratives corroborates and extends previous research conducted with adults (e.g., Dong, Mak, et al., 2024). This makes sense, given that more generally, reading is a crucial avenue for vocabulary acquisition. Previous corpus research shows that as compared to aged-matched spoken language, book language is more emotionally arousing and books use a greater diversity of emotion words (Dawson et al., 2021; Dong & Nation, 2024). Our results also revealed an emotion advantage in *speeded recognition*. Words that were embedded in more emotional contexts – both positive and negative – were recognised more accurately than words presented in neutral contexts. This supports the affective embodiment account, which posits that emotional content enhances cognitive processing and memory retention. Additionally, we observed an age effect in *speeded recognition*, with older children outperforming younger children. This age effect could be attributed to a combination of factors, including increased proficiency in language and cognitive skills. Older children are more likely to have better language and reading, which in turn can support new word learning.

Children were also able to infer word valence from context. In *valence judgment*, performance was above chance, though it is worth noting that performance was relatively low, and the interpretation should take into consideration of the overall low performance. The task was challenging as it did not provide any narrative context during the test. This categorical judgment to either neutral, negative, or positive may be particularly complex for children, who might not have a fully developed understanding of these valence terms or may interpret them differently compared to adults. Complementary evidence was provided by performance on *meaning generation*. Here, children were asked to provide an English word that they thought best represented the meaning of each novel word in context. We did not provide a specific definition of the novel word, so we did not judge the accuracy of their responses in terms of the precision. Instead, we focused on the valence of the English words they produced. There was a close correspondence between the emotional valence of the generated meanings and the relative emotional valence of the context in which the novel words appeared.

General Discussion

Through two studies, we showed that children's writing contains reliable patterns of association between word valence and the valence of the contexts that the word appears in, and that narrative valence is an informative cue for word learning. This might be especially important for learning emotional and valenced words as they tend to be more abstract, and more context-dependent (Davies et al., 2020; Dawson et al., 2021). Therefore, narrative context might enhance children's understanding of emotional concepts. Analysis of infant-directed speech (Nencheva et al., 2023) shows that the language used by caregivers not only communicates information but also reflects the emotional valence of the situation being described. For instance, when describing joyful events, caregivers tend to use positively charged words. This alignment between language and emotional context helps infants not only understand the meaning of the words but also connect those words to the emotional atmosphere. This creates a rich language environment that can support children's emotional development and their ability to associate emotions with specific words and contexts. Our study extends the importance of context for emotional language to older children in the context of written language.

Many studies have explored the influence of context on word learning (e.g., Bolger et al., 2008; Hulme et al., 2023; Mak et al., 2021; Rosa et al., 2022), but relatively few have examined the affective aspects of word learning in adults (e.g., Snefjella et al., 2020), and none have focused on children, especially through natural reading. Our findings suggest that children as young as 7 years old can infer the emotional valence and meaning of novel words through reading in context, even without explicit definitions or scaffolding by adults. This highlights that word learning occurs incidentally and incrementally during reading. In our study, although we asked children to pay attention to the novel words, they were not informed about any upcoming tests. While we acknowledge that our approach might not fully resemble natural reading, where children are not typically instructed to focus on unfamiliar words, our approach is a step closer to incidental learning. Future research could eliminate this instruction entirely to assess whether children can infer valence

spontaneously. Moreover, in line with our aim to simulate incidental learning, we chose not to provide predefined meanings for the made-up words. This allowed children to engage with the novel words more naturally. Consequently, we did not control for whether the novel adjectives were used in prenominal (before nouns) or postnominal (predicative) positions. Yet, real words sometimes have positional constraints (e.g., Davies et al., 2020); for example, “afraid” can only be used predicatively. Future studies should consider controlling for adjective position and explore the syntactic, semantic, and pragmatic dimensions of adjectives to provide a more comprehensive understanding of children's acquisition of adjectives via reading.

Our results support the affective embodiment account that emotional valence facilitates the acquisition of abstract words, but the account does not specify whether positive or negative valence is more effective in supporting word learning and processing. Indeed, our exploratory analysis showed no significant difference between positive or negative contexts in *speeded recognition*. Interestingly, however, when looking at children’s own language production, both children’s writing in Study 1 and children’s generated meaning in Study 2 showed a linguistic positivity bias. This is in line with the Pollyanna hypothesis that human language demonstrates a universal tendency to use more positive terms (Boucher & Osgood, 1969; Dodds et al., 2015). In *meaning generation* in Study 2, the overall valence of all generated words is 5.49 out of 9, which is also above the midpoint of 5. Specifically, the estimated valence rating for words in the neutral context (5.90) is also slightly more positive than neutral. Similar findings were also reported in adult participants in Dong, Mak, et al. (2024). In Study 1, and in adult language, as reported by Snefjella and Kuperman (2016), the observation that mean word valence and context valence are both higher than the midpoint of 5 shows greater usage of more positive words and contexts. These provide additional evidence for the Pollyanna hypothesis and the universal linguistic positivity. While both positive and negative contexts support learning, when it comes to language usage, both adults and children tend to use more positive language.

Additionally, while the better performance in emotional contexts is in line with the affective embodiment account, it does not preclude alternative explanations. Broader language learning

mechanisms, such as statistical learning and distributional semantics may also contribute to children's learning of emotion language (Lenci et al., 2018; Seidenberg & MacDonald, 2018). To rule out other possible explanations such as word association in accounting for word learning, future studies should more strictly control for stimulus attributes, including arousal and dominance. Additionally, future studies should incorporate tasks that assess children's understanding of novel word meanings beyond valence, such as testing their ability to use the novel words correctly in new, unrelated contexts. These approaches would help determine whether it is the emotional contexts that facilitate deeper semantic learning rather than simple associative links.

One concern about Study 1 might be that the analyses used ratings of valence, arousal, and concreteness from adult English speaker norms. This was necessary as there are no equivalent large datasets of ratings by children. Future work should look to establish children's ratings and compare them with those of adults, but meanwhile, we note that the pattern of correlations of adult ratings across children's writing resembled that of adult ratings across adult language usage, suggesting commonality across age. Additionally, English was the language used in all the studies referenced in this paper, and it would be interesting to explore whether similar patterns might be observed in other languages, both for adults and children. Future research could provide a cross-cultural perspective to determine the extent to which the effects reported here are specific to English or are observed across different linguistic and cultural contexts.

In conclusion, we combined a data-driven and an experimental approach to understanding emotional context in children's language. We first demonstrated through analysing children's own writing that context valence correlated with word valence. We then tested through an experiment that children were able to infer the emotional valence of novel words from reading them in context, and that words read in emotional contexts were learned better than those experienced in more neutral contexts. As the first study of its kind with children, questions remain for future research to address. For example, assessing word knowledge after a delay would reveal whether the learning effects seen in our experiment maintain over time, and introducing a frequency manipulation could test how learning

unfolds through multiple exposures. We focused on valence, and it would be interesting to consider other emotion attributes such as arousal. Together, this will deepen our understanding of the interplay between words and the contexts in which they appear, and how affective associations of both influence word learning.

In the same data collection session as reported in this Chapter, I also included a few other tasks, which allows me to explore in the next chapter how children's experience with reading and written language might be related to the emotional content of the language they produce. This would link with corpus findings in Chapters 2 and 3, highlighting the importance of reading.

Chapter 6. Reading Experience and Emotion Language Production

Chapter summary:

This chapter takes an exploratory step towards understanding how the language children encounter through reading influences their verbal and emotion language production. Previous corpus analyses of emotion language in children's reading and writing have shown similar patterns that language written for and by older children exhibit greater diversity and complexity in emotional terms compared to younger children. This chapter bridges these findings by directly examining the relationship between individual differences in exposure to book language and children's emotion language production. Children's exposure to book language was assessed by an author recognition test and word reading efficiency, and their production of emotion words was measured both in isolation and in narrative contexts. Results showed that the lexical diversity in children's spoken production correlated with their exposure to book language, reinforcing the corpus findings presented in Chapter 2. Children who read more also tended to enjoy reading more, or vice versa. However, while positive correlations were found between print exposure and production of emotion words in narratives, print exposure was not a significant predictor once total word count was controlled for.

Introduction

Extensive research indicates that reading experience may enhance linguistic skills and contribute to social-emotional development (Batini et al., 2021; Mol & Bus, 2011). The constructionist approach to emotional development suggests that language is a key component of emotional growth, supporting the development of various aspects of emotional understanding (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017; Shablack & Lindquist, 2019). Relatedly, Nation et al. (2022) argued that exposure to written language provides children with opportunities to experience language that is different from everyday usage, including the complex and nuanced language associated with emotions and mental states. They suggested that the language needed to drive emotional development through childhood might build directly from experience with written language, through reading or being read to. To test

this idea further, it is important to understand how the language children encounter through reading influences their verbal and emotion language abilities. Chapters 2 and 3 reported corpus research that separately examined the emotion language in children's reading and writing (Dong, Hsiao, et al., 2024; Dong & Nation, 2024). Both chapters have shown similar patterns that language written for and by older children exhibit greater diversity and complexity in emotional terms compared to younger children. However, a direct connection between individual differences in exposure to book language and the emotion language produced by a child has not been explored. This final chapter is an exploratory first step to attempt to integrate the different approaches and themes together, trying to connect children's reading experience to their emotion language production.

Emotion understanding is closely associated with language competence. For example, early language ability is a strong predictor of children's later ability to recognise emotional cues (e.g., Beck et al., 2012; S. Griffiths et al., 2020; Widen & Russell, 2003, 2008). The ability to use language to identify and label emotions, both in oneself and in others, has been described as essential for successful social interactions and mental health. S. Griffiths et al. (2020) found that children with more advanced language skills at the ages of 5 to 6 years were significantly better at identifying emotions from facial and vocal cues by 10 to 12 years old, whereas children with developmental language disorder faced notable challenges in recognising emotions. Research with school-aged children has demonstrated a broader connection between language and emotional competence. Beck et al. (2012) found that receptive vocabulary and literacy were closely linked to emotional knowledge, particularly the ability to recognise mixed emotions. In summary, there is evidence to show that as children grow and encounter increasingly complex social situations, language skills are implicated in navigating emotions.

One way that language supports emotional development is through the acquisition of emotional vocabulary. Reading experience and exposure to book language, also known as print exposure, might play a key role in the development of emotional vocabulary. Reading provides children with access to diverse emotional experiences and to vocabulary that is less frequently encountered in everyday

conversation (Dawson et al., 2021; Nation et al., 2022). Focusing on emotion and mental state language more specifically, my cross-corpus analyses reported in Chapter 2 show that books contain a broader range of emotional and mental-state words than child-directed speech or child-directed television programs (Dong & Nation, 2024). This opportunity to engage with complex and diverse emotion language may allow children to experience emotions they may not otherwise encounter, broadening their emotional knowledge. Certainly, there are significant variations in independent reading habits among children, with some reading daily and others only sporadically (e.g., Mol & Bus, 2011; Van Steensel, 2006). It is plausible that individual differences in print exposure might lead to disparities in emotional knowledge, where children who read more frequently may be better equipped with emotional vocabulary and have a more nuanced understanding of emotions. Relatedly, print exposure has been linked to the Theory of Mind (ToM), which involves understanding others' mental and emotional states. There is a positive correlation between print exposure and performance on ToM tasks in adults leading to the suggestion that reading, particularly narrative fiction, is closely tied to emotional and cognitive empathy (e.g., Kidd & Castano, 2013; Mar et al., 2006). In children, Adrian et al. (2005) found that children aged 4 to 5 years who experienced more parent-child storybook reading at home experienced greater use of mental state terms by their mothers during shared reading, and they performed better on false belief tasks. This leads to the speculation that as children get older and have the opportunity to experience complex characters and emotionally rich stories via reading, independent reading might further allow them to practice understanding emotions and perspectives different from their own.

To examine children's reading experience, it is important to have reliable measures to assess children's exposure to book language. The Author Recognition Test (ART) has been widely used as a proxy for print exposure (Brysbaert et al., 2020). In the ART, the names of popular fiction authors are mixed with the names of unknown individuals, and participants are asked to indicate which authors they know. It was first developed by Stanovich and West (1989) and subsequently adapted and validated by numerous other researchers (e.g., Acheson et al., 2008; Brysbaert et al., 2020; Fong et al., 2013; Mar & Rain, 2015). Performance on the ART has been shown to correlate strongly with

vocabulary and reading ability, as well as emotional development, including ToM and empathy (e.g., Kidd & Castano, 2013, 2019). However, existing ARTs for children in the UK are outdated and include authors that might no longer be relevant to today's young readers (e.g., Ricketts et al., 2007; Stainthorp, 1997). To address this, I developed an updated ART specifically tailored to primary school children in the UK with the aim of providing a more accurate measure of children's print exposure today.

Building on this foundation, this chapter sought to explore how children's print exposure and language skills relate to their production and use of emotion language. Use of emotional-state talk has been found to correlate with 7-to-10-year-old children's emotion understanding (Grazzani & Ornaghi, 2012; Ornaghi & Grazzani, 2013). Previous research has examined children's production of emotion words for the purpose of constructing an emotion lexicon (e.g., Doost et al., 1999; Ridgeway et al., 1985). For example, Doost et al. (1999) used a word generation task which asked children to produce as many emotion words as they could to describe a particular feeling in an imagined emotional scene (e.g., "think about a child who is feeling sad and then write down as many words as you can to describe this feeling"). While these studies have examined factors that influence the emotion lexicon (e.g., age), they have not considered whether and how reading experience might contribute to emotion word production.

In this chapter, production is measured through two complementary tasks: verbal fluency and narrative production; the former capturing emotion word production in isolation, and the latter examining its use in broader narrative contexts.

Emotion fluency measures the ability to produce emotion words under time constraints (Abeare et al., 2017, 2022). Several studies have measured emotion fluency in adults, but the results have been mixed regarding whether emotion fluency is related to emotional well-being. Camodeca et al. (2021) linked emotion fluency with social cognition, such as ToM, indicating that generating emotion words may also tap into social-emotional competence. However, Hegefeld et al. (2023) found that emotion fluency was related to general verbal fluency but did not directly predict emotional well-being. To my

knowledge, only one study has looked at the emergence of emotion fluency though development. In a cross-sectional sample of participants aged 5 to 25 years, Valdivia-Moreno et al. (2023) found that both emotion fluency and verbal fluency showed similar increases across age and are highly correlated, plateauing in late adolescence. Once again, however, the potential role of reading has not been considered. Thus, the work reported in this chapter aimed to document emotion fluency in school-aged children and examine whether this is correlated with other factors, including reading experience (as measured by the ART) and measures of reading efficiency. This will allow me to further explore the relationship between print exposure and emotional vocabulary.

To complement the emotion fluency task, I developed a narrative production task that examined how children use emotion language in a storytelling context. Previous research has shown that older children's written narratives tend to display greater lexical sophistication and diversity than those of younger children (Hsiao et al., 2024), and that boys and older children produced less positive emotional content (Dong, Hsiao, et al., 2024). By analysing children's narratives, I aimed to understand how emotional vocabulary is represented in their narrative production, and how this is associated with print exposure.

This chapter is an initial step towards integrating the theme from language experience to language production, as alluded to throughout this thesis. To summarise, I measured children's print exposure through an updated ART and investigated its relation to both emotion language and general language production. Specifically, I explored whether print exposure and reading ability predicted the lexical diversity and the diversity of emotion words produced in their language, as indexed by their performance on the fluency tasks as well as narrative production. Finally, and linking back to the approach used in Chapter 3, I also explored the sentiments of the children's narrative production and examined their association with age and gender.

I hypothesised that reading experience might contribute to greater diversity in language production by exposing children to greater lexical richness in books. Based on this hypothesis, I predicted that individual differences in children's reading experience and reading skills would be associated with

variation in the (a) lexical diversity of their language production, and (b) diversity of emotion words in their language production. Additionally, following findings discussed in Chapter 3, the sentiment of children's narrative production would be associated with children's age and gender, with older children and boys producing less positive content.

The study hypotheses, measures and the preliminary analyses were preregistered: https://doi.org/10.17605/OSF.IO/6MWDZ. Analysis codes and materials are available on OSF (http://tiny.cc/rainy_thesis_chapter6).

Method

Participants

The participants were the same group of children who participated in the experiments in Chapter 5, recruited from primary schools around the UK. Two children were excluded from the original 120 datasets collected as these two children did not pass the preregistered exclusion criteria as set out in Chapter 5. After exclusion, datasets from 118 children (57 girls, 61 boys) were retained for analysis. The ages of the 118 children ranged from 7 to 11 years (M_{age} = 9.71, SD_{age} = 1.12). All children were enrolled in primary school in the UK. Fifty-four children (45.8%) had at least one family member who spoke to them in a language other than English. Informed consent was received from parents or guardians, and children provided verbal assent before starting the experiment.

Materials

Children completed a battery of tests, including existing standardised measures as well as tasks created for the purpose of the study.

Test of Word Reading Efficiency (TOWRE)

I administered the two subtests of the Test of Word Reading Efficiency Second Edition (TOWRE–2; Torgesen et al., 2012) to measure children's word reading fluency. The sight word subtest measured

how many English words (up to a total number of 104), ranging from high to low frequency of occurrence, students could accurately read aloud in 45 seconds. The phonemic decoding subtest assessed how many nonwords (up to a total number of 63), varying in complexity, students could accurately pronounce in 45 seconds.

Author Recognition Task (ART)

I first compiled a list of children's authors based on an author list by Ricketts et al. (2007), adding in more recent authors of the Popular Children's books in Waterstones website (https://www.waterstones.com/category/childrens-teenage/, accessed in early 2024). After consulting with several parents of primary school children and primary school librarians, as well as an informal pilot, I shortened the list of authors to 30, as I wanted the test to be relatively brief. Following Brysbaert et al. (2020), I opted for a higher number of authors and a lower number of non-author foils, to ensure good sensitivity and reliability of the test to detect potential variability among children's print exposure. The final list contained 50 items, 30 author names and 20 non-author foils. See Appendix 6A for the complete list of items.

Verbal Fluency

The verbal fluency test taps children's vocabulary knowledge of two categories: animals and emotions. Children were asked to produce as many animal names or emotion words as they can in 30 seconds.

Narrative Production

Children were asked to narrate a story about a memorable day in their life within a two-minute limit. Given their young age, and to prevent developing skills like spelling and handwriting from hindering their ability to express themselves, I opted for a spoken narrative production instead of a written one.

Procedure

All measures were administered by a trained member of the research team in a quiet room adjacent to the main classroom. Assessors underwent comprehensive training, which included the completion of online safeguarding training modules (including quizzes) and two in-person training sessions. A protocol was also provided with detailed procedures and instructions to ensure consistent measurement administration across schools. For all tasks that involved verbal production, children's responses were audio-recorded and subsequently transcribed for checking and further processing.

After introducing the context of the study which involves an alien friend, children first completed the sight word reading subtest of TOWRE. Following that, children completed the verbal fluency measures, including the animal fluency task followed by the emotion fluency task in the same fixed order. Children were asked: "Can you tell me as many animal names (or emotion words) as you can in 30 seconds?" They then completed the word learning task reported in Chapter 5, followed by the phonemic decoding subtest of TOWRE.

The ART was administered online via Gorilla (www.gorilla.sc, Anwyl-Irvine et al., 2020) and it took around 5 minutes to complete. Each trial began with a fixation cross displayed for 250 ms, followed by an author name or foil and the child was to press the tick or the cross mark on the keyboard to make a judgment. After a practice trial, children responded to the 50 items (30 author names and 20 foils), presented one at a time in random order in a single block. No feedback was provided and there was no time constraint. After completing the ART, children were asked about their reading enjoyment: "How much do you enjoy reading?". They answered with a number by dragging a slider that ranged from 1 (not at all) to 100 (very much).

Finally, children completed the narrative production task where the alien friend asked them to "tell me about a memorable day in your life". They were given three guiding questions to get them started: "Where did you go? What did you do? How did you feel?", and they had 2 minutes to tell their story. They could stop at any time but were encouraged to say more with gentle and neutral prompts such as "tell me more about it". At the end of the study, children provided basic demographic information such as age, gender, whether they had at least one family member who spoke to them in a language

other than English. Stickers related to the theme of the study were given as motivation throughout the study.

Data processing and analysis

Standard scores for the TOWRE performance were computed from the raw scores and age, following the instruction manual. For the ART, I followed the guidelines of Brysbaert et al. (2020), calculating the score as the percentage of correctly identified author names minus the percentage of non-authors incorrectly selected. For example, a child who indicated that they knew 18 of the 30 authors and wrongly identified 8 of the 20 foils as authors would receive a score of 18/30 - 8/20 = 0.20, or 20%.

Words produced in the two fluency tasks were transcribed from the audio recordings. All mentions of animals were converted to singular forms. To score these, a reference list of animal names (or emotion words) was compiled from a variety of sources, including journal articles, websites, and corpora (animal names N = 2870, emotion words N = 1198, see Appendix 1 and Appendix 6 for more details). Automatic scoring procedures were developed such that if a child's response matched an item on the reference list, they received a score of one. Items not found on the respective reference list were flagged and manually checked to determine if they fitted the category. If a child produced multiple forms of the same emotion word, e.g., anxious and anxiety, only one instance was counted. The final score was based on the number of words produced that fit within each category.

For the transcripts of children's narrative production, I included all relevant speech by the child, whether spontaneous or prompted by the experimenter. This approach was necessary because most children produced minimal speech independently. In processing the transcripts, I removed inaudible utterances, fillers (e.g., "um", "uh," "yeah"), and comments deemed unrelated to the narrative content, as well as any speech by the researcher. Words such as "well" and "like" were retained, as they could function both as fillers and content words, making them difficult to classify automatically. I also added a space between contracted forms (e.g., replacing "I'd" with "I 'd") to ensure accurate word counts. Each narrative was then segmented into utterances, following the procedures outlined in the

Expression, Reception, and Recall of Narrative Instrument manual (ERRNI; Bishop, 2004). I first examined children's overall language quality, where I calculated the mean length of utterance (MLU) for each child's narrative and, after excluding function words, computed the type-token ratio (TTR).

For emotion-related indices, I calculated the number of affective words in the narratives that had an emotional association, based on existing norms (Warriner et al., 2013). Emotional valence was rated on a scale from 1 to 9, with 1 being very negative and 9 very positive. Words with a valence rating above 6 or below 4 were considered emotional. I also calculated the sentiment scores for each child's narrative using a *Bidirectional Encoder Representations from Transformers* (*BERT*; Devlin et al., 2019) model, following the procedures in Chapters 3 to 5. *BERT* evaluates the context of each word to generate word embedding representations. The sentiment score was a continuous value ranging from -1 (very negative) to 1 (very positive).

For analysis, I started by calculating correlations between the various variables. I then conducted several multiple regressions, to understand how production outcomes might be related to verbal skills and print exposure.

Results

The results section is organised as follows: descriptive statistics, correlation analyses, and planned regression analyses. This chapter reports these confirmatory findings only, due to time constraints. In the future, I plan to conduct further analyses to better understand the relationship between the variables, including structured equation modelling.

Descriptive statistics

Summary statistics are presented in Table 6-1. One child was further excluded as they could not generate any response in the narrative production task ("I can't think of anything"), leaving 117 sets of complete data for all analysis. Histograms depicting the distribution of performance on each task are shown in Figure 6-1. Not all variables were normally distributed. Shapiro-Wilk test for normality

was significant for eight of the twelve variables (these variables are underlined in Table 6-1).

Children on average read 69.09 words (*SD* = 10.17, range 41-104) in the TOWRE sight word reading subtest and 40.73 made-up words (*SD* = 10.67, range 19-62) in the phonemic decoding subtest. As reported in Chapter 5, all children included in this chapter scored above 1.5 standard deviation below the standard mean. Hence, they are judged to be able to participate in the task.

The mean score for ART was 0.35 (*SD* = 0.19). The percentage of authors selected was 49% and foils selected 15%. Internal item consistency was 0.82, measured with split-half reliability by randomly splitting the test items in half. Cronbach's alpha was 0.86. On average, children rated their reading enjoyment as 74.4/100.

For the two fluency tasks, children produced an average of 14 animal names and 8 emotion words in 30 seconds. Figure 6-2 shows a word cloud of the 50 most frequently produced (a) animal names and (b) emotion words. The larger the size of the word, the more often children mentioned it.

In the narrative task, children produced on average 103 words, in which 28 words were emotional words (defined as words with a valence rating above 6 or below 4 in the valence norm that ranged from 1 to 9). The mean length of utterance was 8.93, and the type-token ratio was 0.79. The average sentiment of children's narratives was 0.31, which was positive.

Table 6-1. Summary statistics of measured variables (N = 117).

Task	Variables	Mean	SD	Min	Max	Range	Skew	Kurtosis	SE
	Age*	9.69	1.11	7.00	11.83	4.83	-0.04	-0.85	0.10
TOWRE	TOWRE Sight Word	69.09	10.17	41.00	104.00	63.00	-0.06	0.39	0.94
	TOWRE Nonword **	40.73	10.67	19.00	62.00	43.00	-0.35	-0.77	0.99
ART	ART Score	0.35	0.19	-0.30	0.78	1.08	-0.05	0.33	0.02
	Enjoyment***	74.44	22.73	0.00	100.00	100.00	-1.00	0.87	2.10
Verbal Fluency	Animal Fluency	13.66	4.16	4.00	25.00	21.00	0.16	-0.08	0.38
	Emotion Fluency***	8.20	2.71	3.00	17.00	14.00	0.69	0.43	0.25
Narrative Production	Total Word Count***	102.88	74.24	8.00	349.00	341.00	1.12	0.54	6.86
	MLU	8.93	2.26	2.00	16.50	14.50	0.07	0.70	0.21
	TTR*	0.79	0.11	0.59	1.00	0.41	0.14	-0.91	0.01
	Emotional Word Count***	28.34	18.84	3.00	98.00	95.00	1.19	1.11	1.74
	BERT Score**	0.31	0.36	-0.56	0.89	1.45	-0.29	-0.84	0.03

Note: variables that are not normally distributed are underlined. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$ for the associated p-values for the Shapiro-Wilk Test for Normality.

Figure 6-1. Histograms showing the distribution of the 12 variables.

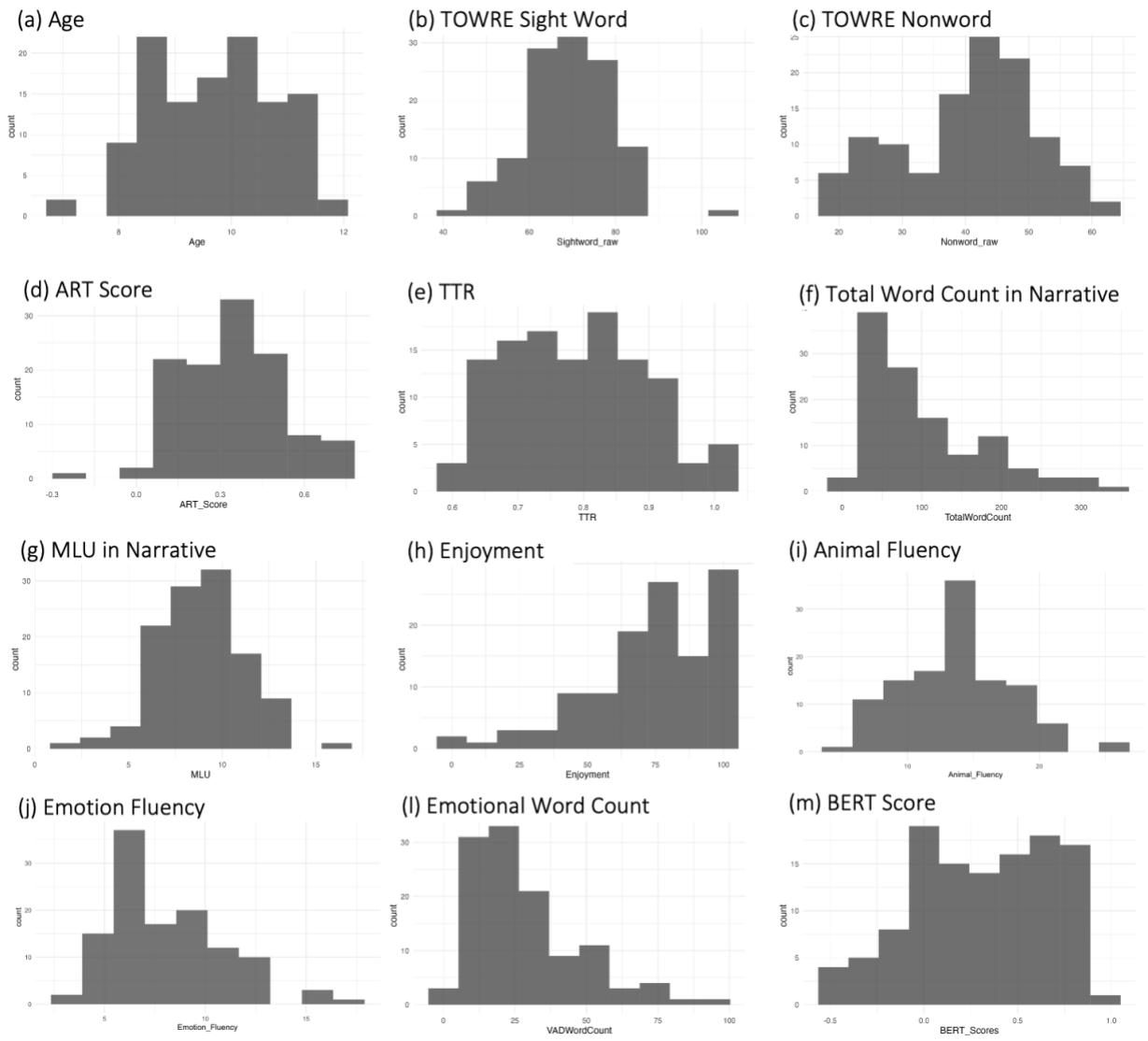
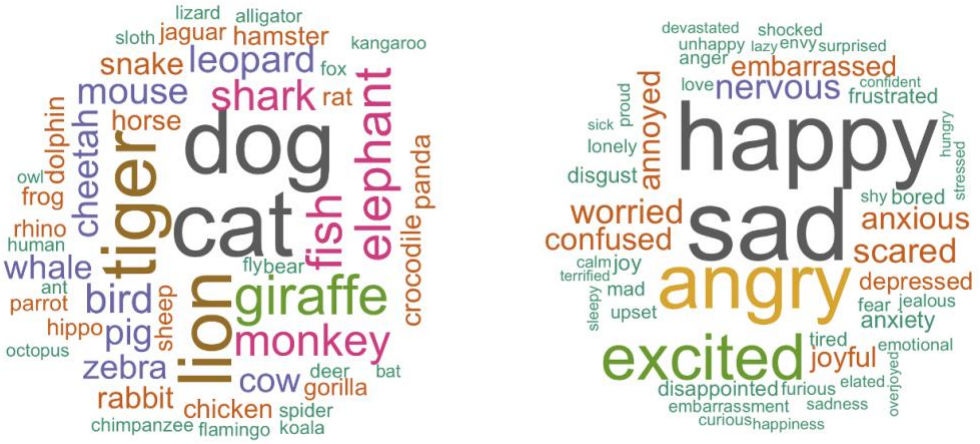


Figure 6-2. Word cloud of 50 most frequently produced (a) animal names and (b) emotion words in the two versions of the fluency task.

(a) Animal fluency (b) Emotion fluency



Note. Multiple forms of the same emotion words (e.g., anxiety and anxious) were shown in the word cloud to represent the range of emotion words children produced. However, they were counted as a single emotion when scoring for emotion fluency.

Correlational analyses

Table 6-2 presents the correlations between performance on the different tasks. Since most variables were not normally distributed (underlined in Table 6-1), I used Spearman's correlation to maintain consistency and comparability of the results. Age was positively correlated with all variables except for enjoyment and narrative sentiment (no significant correlation) and type-token ratio (negative correlation), indicating that older children tended to have better language skills and produce more emotion words. Additionally, ART scores were positively correlated with enjoyment, suggesting that children who enjoyed reading more tended to read more, or vice versa. ART scores were also positively correlated with TOWRE sight word and nonword reading, animal fluency, total word count and emotional word count in narrative production. This suggests that children who read more also

tended to have better word reading ability and verbal fluency, and to produce longer narrative that included more emotional words. Emotion fluency correlated positively with animal fluency, but not with ART scores. Type-token ratio correlated negatively with age, total word count, and emotional word count. This shows that children who were older, produced longer narratives and included more emotional words in their production might also be likely to have smaller lexical diversity in their narrative production. The sentiment of children's narrative production correlated negatively with the total word count and the emotional word count, suggesting that longer narratives tended to be less positive and to include more emotional words. These associations are then further explored in the regression analysis.

Table 6-2. Correlation matrix of the variables.

	Age	TOWRE Sight Word	TOWRE Nonword	ART Score	Enjoy-ment	Animal Fluency	Emotion Fluency	Total Word Count	MLU	TTR	Emotional Word Count	BERT Scores
Age	1.00											
TOWRE Sight Word	0.35***	1.00										
TOWRE Nonword	0.27*	0.73***	1.00									
ART Score	0.44***	0.45***	0.55***	1.00								
Enjoyment	-0.04	0.19	0.19	0.21*	1.00							
Animal Fluency	0.37***	0.30**	0.26**	0.50***	0.20*	1.00						
Emotion Fluency	0.26**	0.21*	0.12	0.20*	0.18*	0.41***	1.00					
Total Word Count	0.29*	0.33**	0.27**	0.22*	0.16	0.23*	0.18	1.00				
MLU	0.20*	0.08	0.03	0.17	0.20*	0.32***	0.23*	0.49***	1.00			
TTR	-0.25*	-0.18	-0.12	0.03	0.03	-0.02	0.04	-0.61***	-0.15	1.00		
Emotional Word Count	0.33**	0.34***	0.28**	0.24*	0.20	0.23*	0.14	0.96***	0.42***	-0.65***	1.00	
BERT Score	-0.01	0.01	-0.02	0.10	0.06	-0.03	-0.16	-0.28**	-0.05	0.16	-0.23*	1.00

Note: *p*-values were adjusted for multiple comparisons using False Discovery Rate (FDR) correction, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Regression analyses

As the variables were on different scales, I first standardised them to allow for meaningful interpretations and comparisons. I then ran multiple regression models based on the hypotheses. Table 6-3 presents all model outputs.

Quality of language production

I started by analysing whether the quality of children's language production is associated with their reading experience and ability. The quality of children's language production was assessed with the mean length of utterance and type-token ratio of narrative production. Children's age, print exposure, TOWRE sight word and nonword reading, and the total word count in their narrative production explained 18% of variance in mean length of utterance (Model 1a), $F(5,111) = 4.94$, $p < .001$, and the only significant predictor was total word count.

For type-token ratio (Model 1b), the four variables explained 42% of variance, $F(5,111) = 16.21$, $p < .001$. Age, ART, and total word count were significant predictors of type-token ratio. Younger children who read more, and those who produced less content showed greater lexical diversity in their narrative production. Note that for both mean length of utterance and type-token ratio, TOWRE sight word and nonword reading did not explain unique variances. This suggests that the word reading efficiency did not predict quality of language production in this samples.

Production of emotion words

Zooming into emotion language, I also analysed whether production of emotion words in the emotion fluency task and in the narrative production task was associated with print exposure and reading ability. Children's age, print exposure, word and nonword reading, and animal fluency explained 16% of the variance in emotion fluency (Model 2a), $F(5,111) = 7.18$, $p < .001$. Word reading and animal fluency were significant predictors of emotion fluency. Children who were better word readers and produced more animal names also tended to produce more emotion words.

For emotional word count in narrative production (Model 2b), children's age, print exposure, word and nonword reading, and total word count explained 93% of variance, $F(5,111) = 302.3$, $p < .001$. Total word count was the only significant predictor.

Sentiment in children's language production

Finally, I analysed whether the sentiment of children's narratives is related to their age and gender, controlling for word count. Gender was a significant predictor for narrative sentiment, where narratives produced by girls were more positive than those produced by boys. Total word count was also a significant predictor of narrative sentiment, where longer narratives tended to have less positive sentiments.

Table 6-3. Regression model outputs (N = 117).

Model	Predictor	Coefficient (SE)	R^2	Adj_R^2
Model 1a: Outcome - MLU			0.18	0.15
	(Intercept)	0 (0.09)		
	Age	0.12 (0.10)		
	ART Score	0.12 (0.11)		
	TOWRE Sight Word	-0.02 (0.14)		
	TOWRE Nonword	-0.15 (0.14)		
	Total Word Count	-0.37 (0.09) ***		
Model 1b: Outcome - TTR			0.42	0.40
	(Intercept)	0 (0.07)		
	Age	-0.24 (0.09)**		
	ART Score	0.30 (0.11)**		
	TOWRE Sight Word	0.01 (0.11)		

TOWRE Nonword	-0.01 (0.12)		
Total Word Count	-0.57 (0.08)***		
Model 2a: Outcome - Emotion Fluency		0.24	0.21
(Intercept)	0 (0.09)		
Age	0.14 (0.09)		
ART Score	0.02 (0.11)		
TOWRE Sight Word	0.31 (0.13)*		
TOWRE Nonword	-0.19 (0.14)		
Animal Fluency	0.33 (0.10)***		
Model 2b: Outcome - Emotional Word Count		0.93	0.93
(Intercept)	0 (0.02)		
Age	0.04 (0.03)		
ART Score	0.03 (0.03)		
TOWRE Sight Word	-0.00 (0.04)		
TOWRE Nonword	0.01 (0.04)		
Total Word Count	0.95 (0.03)***		
Model 3: Outcome - Narrative Sentiment		NA	NA
(Intercept)	0.21 (0.04)***		
Age	0.04 (0.03)		
Gender^	0.21 (0.06)**		
Total Word Count	-0.11 (0.03)***		

Note. *$p < .05$, **$p < .01$, ***$p < .001$. ^Gender is a binary predictor with 0 as boy and 1 as girl.

Discussion

This chapter examined whether children's reading experience is related to their language production,

and in particular production of emotion words. I first developed an updated ART as a proxy for children's print exposure. This had good internal validity and correlated well with age and reading ability. The positive correlation between print exposure and reading ability aligns with previous studies showing that print exposure and reading skills are strongly associated across development (Mol & Bus, 2011). I also found a small but significant positive correlation between ART score and reading enjoyment, suggesting that children who read more tend to enjoy reading more, or vice versa. While the results are correlational, a study by van Bergen et al. (2023) tested direction-of-causation in 3690 12-year-old twins. Their analyses suggested that children with good literacy are more likely to enjoy reading and to read more. This finding highlights the importance of fostering children's reading ability. It is important to note that this newly developed ART included only 30 author names, carefully selected after an initial pilot to capture variability while keeping the task brief and easy to administer. However, I recognise that many other authors whose books the children have read might not have been included. Future studies could explore alternative measures, such as an author fluency task, where children name authors they know within a time limit (McCarron et al., 2024). This approach might better capture children's knowledge of authors while engaging them in a manner more suited to their cognitive abilities. Additionally, investigating whether fluency-based tasks and recognition tasks tap into different dimensions of print exposure could provide complementary insights.

Regarding the first prediction on the quality of narrative production, the two measures of type-token ratio and mean length of utterance revealed different patterns. For type-token ratio, age, print exposure, and total word count were significant predictors. This indicates that print exposure explains unique variance in type-token ratio, even after controlling for age and total word count. Since book language is more decontextualised and lexically richer, it is plausible that children who read more show greater lexical diversity in their speech. The results support the findings by Hamilton et al. (2020), which found that measures of fiction reading experience correlated positively with indices of narrative production quality and predicted unique variance in narrative macrostructure. It is worth noting that both age and total word count predicted type-token ratio in the negative direction. It makes sense for total word count, as longer texts tend to reuse a core vocabulary, reducing lexical diversity

relative to text length (Richards, 1987). Future research should consider using alternative measures of lexical diversity that account for text length such as moving-average type-token ratio (e.g., Fergadiotis et al., 2015; Yang et al., 2022). The negative correlation between age and type-token ratio was more unexpected, which contradicts previous work showing the opposite pattern (Hsiao et al., 2024). One of the explanations might be that older children tended to produce more content, which was then associated with lower type-token ratio. Mediation analysis could further test the relationship.

Mean length of utterance, on the other hand, was only strongly predicted by total word count, but not well-explained by these other predictors. This is not entirely surprising, as children's narratives often contained incomplete sentences, making it difficult to establish clear utterance boundaries. Children's inconsistent use of filler words and repeated words further complicated the analysis. Although I excluded certain fillers such as "uh", I retained others such as "like" and "well" because they can serve meaningful roles in discourse, making it hard to judge in cases of incomplete utterances. Similarly, I did not remove repeated words, which might reflect hesitancy or emphasis in spoken production. Moreover, the scoring procedures from the ERRNI manual were designed for standardised narrative retelling (Bishop, 2004), where all children listen to the same story and then see the same picture prompts when they retell the story. This is different to the less constrained format of the narrative production task, which might invite greater variability in terms of length and content. It was also challenging to motivate the children to produce sufficient content during narrative production, and children who were more eloquent and confident in language might tend to speak more. It is therefore not surprising that total word count strongly predicted mean length of utterance, as well as type-token ratio. As such, our results partially support the prediction that individual differences in children's reading experience predicted children's lexical diversity in narrative production, but not their mean length of utterance, likely due to the huge diversity in the length of production.

To address the second prediction, production of emotional words in narratives was also only significantly predicted by total word count, similar to mean length of utterance. This makes sense as

the more content children spoke about, the more words would be scored as emotional. While the production of emotional words correlated with children's age, reading experience, they were not significant predictors once total word count was controlled for. Together with the result for mean length of utterance, it might be that neither mean length of utterance nor emotion word count accurately reflect the children's language proficiency. Future research could benefit from refining task prompts to focus on narrower, more specific topics that are more accessible to children, such that there might be less variability in the total length of production.

For the production of emotion words in isolation, analysis of emotion fluency showed a positive correlation between emotion fluency and animal fluency, and that animal fluency remained as a significant predictor in the regression model. This suggests a close matching between the two category fluency tasks, and that emotion fluency is likely a specific instance of verbal fluency, supporting previous work with adults (Hegefeld et al., 2023) and children (Valdivia-Moreno et al., 2023). While age and ART were positively associated with emotion fluency, they were not significant predictors for emotion fluency after controlling for animal fluency and word reading efficiency. Thus, the results do not support the prediction that individual differences in print exposure predicts emotion word production.

To date, I have only counted the total number of emotion words produced in the emotion fluency task and not further categorised the types of emotion words produced, or the order in which they were generated. Valdivia-Moreno et al. (2023) found that participants produced more negative emotion words than positive or neutral words, and these proportions were invariant across the age span (5-25 years in their study). They also found that the accessibility of each emotion word, indexed by the order of production, was associated with lexical properties such as age of acquisition and concreteness. Follow-up analysis will examine children's production of emotion words beyond the total count. Moreover, emotion word comprehension and use of mental state languages might tap on different cognitive abilities (Grazzani & Ornaghi, 2012), and future study should include a comprehension task to test whether increases in ART may better predict comprehension of emotion

words.

For the sentiment of the narratives, I found that the sentiment of the narratives was predicted by gender, which supports previous findings that girls tend to write more positively in their narratives, as reported in Chapter 3 (Dong, Hsiao, et al., 2024). This gender difference in emotional expression is consistent with established research on the development of emotion expression in children (Chaplin & Aldao, 2013). Additionally, I found that total word count was a significant predictor of narrative sentiment, and an increase in the total word count led to a decrease in overall sentiment. A possible explanation is that memorable events that children were willing to tell more about might include more complex and mixed emotions, and ups and downs, rather than entirely positive. Follow up study employing qualitative analysis may help offer complementary insights, especially when the content is too sparse or varied for quantitative measures to capture meaningful trends. Additionally, I did not find any association between sentiment and age, which was contrary to what was found in the large corpus of children's writing in Chapter 3 (Dong, Hsiao, et al., 2024). This might be attributed to the inherent differences between spoken production and written production; as well as the potential lack of power, where the sample size was determined based on experimental manipulation (in Chapter 5) rather than specifically for the multivariate analyses.

Finally, while age was correlated with most variables and was a significant predictor of most outcome measures, it is important to interpret these findings cautiously due to the cross-sectional nature of the study. The variability among children, particularly in a sample collected from seven different schools across the UK, could reflect other factors such as academic abilities or school environments. The cross-sectional design also limits the ability to infer directionality or causality in the observed relationships between print exposure and language production. While the study demonstrates associations, it cannot determine whether print exposure directly influences language development or whether other unmeasured factors (e.g., home literacy environment, cognitive abilities, verbal language skills) are driving the observed effects. Future research should adopt a longitudinal design to track children's print exposure and language development over time, and measure their emotion

language accordingly. In conclusion, this preliminary work is an initial step to integrate the various approaches used in my thesis to examine book language experience, which provides valuable insights into how children's reading experience, gender, and age influence various aspects of language production.

Chapter 7. General Discussion

In this thesis, I studied the emotion words and emotional valence in children's written language, and explored how reading experience is related to emotion language production. I also investigated whether emotional context influences word learning. Across the different chapters, I have explored various aspects of emotion language in context, with each chapter presenting results and discussions specific to its focus. Chapters 2 and 3 utilised a data-driven corpus analysis approach to examine emotion language across different children's language corpora, providing insights into the patterns and usage of emotion words in language written for and by children. To complement this data-driven approach, Chapters 4 and 5 conducted experiments to investigate how emotional context affects word learning by both adults and children. Chapter 6 represents an initial step to integrate corpus and experimental approaches by exploring how language exposure influences language production, with a particular focus on the production of emotion language.

In this final chapter, I synthesise the thesis findings in relation to theories and reflect on the key themes that emerged across the different studies. For each key theme, I evaluate the strengths and limitations and propose directions for future research. The chapter ends with a concluding remark, drawing together the main insights of the thesis.

Revisiting key findings and their theoretical implications

Chapter 2 compared the frequency and diversity of emotion words in children's language experiences, with a focus on written language. It also analysed the language produced by children, revealing an age-related increase in the use of emotion words, mirroring those found in the books written for them. This suggests that the language children produce might be influenced by the emotional content of their reading materials at different ages, supporting the idea that language input and output are closely intertwined. This can be understood through the lens of the Production-Distribution-Comprehension (PDC) framework proposed by MacDonald (2013). The PDC framework states that language production and comprehension are part of a continuous loop, where each process informs and

enhances the other. While originally developed for spoken language, I extended this framework to capture the nuanced approach taken in my thesis, connecting written language input and output.

Chapter 3 studied how emotions might be expressed in children's writing by examining the sentiment of stories written by children. Using the PDC framework, I suggest that children's writing might not only reflect their current language abilities but also shape their future language experiences. As children write with the goal of making their narratives comprehensible to readers, they become more sensitive to the language patterns, supporting future language processing. However, note that Chapters 2 and 3 rely on correlational data, which precludes establishing a causal relationship between reading experience and language production.

Further examining the emotional valence in children's written production, in Study 1 of Chapter 5, I found a correlation between word valence and context valence across a large corpus of children's writing. Adopting the PDC framework, I hypothesised that if children consistently use certain valence properties in their language production, such as positive or negative emotional tones, it suggests an emerging sensitivity to emotional cues. This sensitivity might support novel word learning by providing contextual clues. This hypothesis was tested in the second study in Chapter 5, and the experimental findings demonstrated that emotional contexts better support the learning of new words compared to neutral contexts. Specifically, the valence of children's generated meanings for novel words reflected the valence of the narrative context the novel words were experienced in. This mirrors children's own written language production in the corpus analysis, where word valence positively correlated with context valence.

Chapter 6 explored the reciprocal interactions between experience, comprehension, and production, illustrating how accumulated language experiences might inform and shape language production. The positive correlation between lexical diversity in children's spoken production and their exposure to book language reinforces the corpus findings in Chapter 2. However, due to practical constraint in data collection, spoken language production was tested in Chapter 6, whereas Chapters 2 and 3 focused on written language production in corpus analyses. As mentioned in previous chapters,

written language is more decontextualised and may offer different opportunities for reflection and revision compared to spontaneous spoken language. It is thus possible that individual differences in reading experience are linked to written language production in ways that differ slightly from spoken production. Nevertheless, spoken and written production are closely associated, and early oral narrative skills have been suggested to be provide the foundation for future written narrative skills, potentially serving as a scaffold for transcribing ideas into text (Dockrell & Connelly, 2016; Pinto et al., 2015).

While this thesis applied the PDC framework to link language experience, comprehension, and production, it did not directly test comprehension of emotion words, hence the findings need to be interpreted with this limitation in mind. However, extensive research has shown a strong correlation between reading experience and comprehension (e.g., Cunningham & Stanovich, 1997; Martin-Chang & Gould, 2008). This suggests that despite this limitation, the thesis findings still offer insights into emotion language development and provide a foundation for future research. I propose that by adopting a PDC framework that considers how language is experienced and produced, researchers can gain a more nuanced understanding of how emotional expressions are developed and communicated through language.

Zooming in on the experimental findings in Chapters 4 and 5, I found that adults and children were able to learn new adjectives from reading, and that emotional contexts better support the learning of new words compared to neutral contexts. These results support the affective embodiment account, which argues that abstract words are deeply connected to emotional experiences rather than sensorimotor experiences. It suggests that emotions provide a grounding mechanism for abstract concepts, allowing people to attach meaning to words that do not have a direct physical or sensory counterpart. This perspective challenges the traditional linguistic view, which considers abstract words to be primarily learned and processed through linguistic systems alone (Lenci et al., 2018).

However, the focus of the thesis on the affective embodiment account and showing evidence in support of it does not exclude alternative explanations. Broader language learning mechanisms, such

as statistical learning and distributional semantics may also support children's learning of emotion language (Lenci et al., 2018; Seidenberg & MacDonald, 2018). As discussed earlier within the PDC framework, statistical regularities play a crucial role in language learning, particularly as they accumulate over time (Ruba et al., 2022). When children learn new words, especially abstract ones, they may initially rely heavily on emotional experience to understand their meanings. For example, a child might grasp the concept of "happiness" not through a direct definition but through repeated exposure to the word in positively charged situations. As children's cognitive and linguistic capabilities develop, they may integrate more sophisticated linguistic cues. Overtime, reliance on emotional grounding may decrease, gradually giving way to a more language-based understanding where distributional properties of words, such as frequency and co-occurrence, become more prominent (Vigliocco et al., 2018). This integration of theories allows for a more comprehensive and abstract understanding of language, where emotional and distributional properties work in tandem to support word learning and conceptual development. This highlights the complexity and the importance of a multifaceted approach to understanding language learning in the context of child development.

Rethinking emotion language

What is emotion language?

The concept of emotion language was repeatedly explored and discussed throughout the thesis, highlighting the intricate association between language and emotions. My thesis primarily focused on the lexical level and concentrated on words used to describe emotions in text. Across the different chapters, I navigated between different conceptualisations of emotion language. The different methodological approaches required distinct operational definitions of what constitutes emotion language, its scope, and the methods for its measurement. The two primary approaches I employed were (1) a narrower definition focusing on emotion-descriptive vocabulary based on Baron-Cohen et al. (2010), and (2) a valence-based emotion association framework, which considers the overall emotional direction of language by leveraging existing valence norms. While each approach has its

limitations and does not permit a full exploration of emotion language, engaging with both revealed the potential insights they offer and has deepened my understanding of the nuances inherent in the concept.

In Chapter 1, while reviewing existing literature on the development of emotion term lists for language and emotion research, I noted that different sources employed different definitions of what constitutes an emotion word, and that most did not consider written language. Before commencing the corpus analysis in Chapter 2, I attempted to compile a relatively inclusive, though not exhaustive, list of emotion words by drawing from a variety of sources, including journal articles, norming studies and written language corpora (reported in Appendix 1). This effort resulted in more than a thousand emotion-related words, yet I felt it still fell short of capturing the full scope of emotion language. Rather than delving into the nuanced variations and synonyms that characterise emotional vocabulary while struggling to draw a definitive boundary for my ever-expanding list, I ultimately decided to take a more focused approach. For the cross-corpus comparison presented in Chapter 2, I leveraged a predefined list of emotion words in aggregate, specifically the 336 emotion-descriptive vocabulary items from Baron-Cohen et al. (2010). This “detour” in defining an appropriate scope for emotion language highlights a central challenge of distilling the essence of an emotion from the myriad linguistic expressions that convey it. Interestingly, the emotion word list I initially compiled found practical application later in my thesis. In Chapter 6, I used this list to assess children’s emotion word fluency, where a more extensive inventory was needed to score children’s production of emotion words. The list proved to be exactly what was needed for this purpose.

With “emotion language in context” as the central theme of my thesis, the importance of context in understanding emotion language cannot be overstated. The meaning of emotion words and expressions can shift depending on the context in which they are used. Situational backdrop, cultural norms, and conversational nuances all play crucial roles in the interpretation of emotion words. For instance, the word “sick” traditionally denotes a state of physical illness, but in contemporary colloquial usage, it can also mean “cool” or “impressive”. Recognising the limitations of considering

emotion words in isolation in Chapter 2, I broadened the scope in Chapter 3 by incorporating sentiment analysis to capture all words with emotional associations. In that chapter, I employed multiple approaches to sentiment analysis, ranging from a lexicon-based method utilising existing valence norms to a machine learning approach using a *BERT* model. The machine learning approach, and later developments in large language models (LLMs), demonstrated clear advantages in accounting for context and nuances in language (e.g., Eke et al., 2021; Talukdar & Biswas, 2023). By leveraging computational power and sophisticated algorithms, data-driven techniques make it possible to measure subtle, context-dependent expressions of emotions.

However, even when context is considered, as in Chapter 3, emotion analysis through text is not foolproof, and various factors present challenges and affect its accuracy and reliability. Language is inherently complex, with emotions often conveyed through nuanced elements beyond the text, such as tone of voice, sarcasm, and irony. These subtleties make it challenging for machine learning algorithms to detect and interpret them accurately unless explicitly trained on such variations. For example, sarcasm can invert the apparent sentiment of a statement, leading to potential misclassification by automated systems. A vivid example of this is humorously depicted in *Inside Out 2*, where a genuine praise is dramatically transformed into a sarcastic tone when it passes through the "Sar-Chasm", a scene that never fails to make me laugh. This moment highlights the difficulty of capturing such nuances through the analysis of text alone. It is important to have multimodal approaches that consider a combination of text, tone, facial expressions, and other modalities to accurately capture and interpret emotional tone. Efforts like the Stanford Emotional Narrative Dataset (SEND) by Ong et al. (2021) exemplify an attempt towards capturing these multifaceted expressions, but they also highlight the need for more comprehensive and diverse data to improve model performance. As research progresses, the development of such datasets and models will be critical for advancing our understanding and analysis of emotion language in context.

While LLMs and data-driven approaches shed light on patterns of emotion language through text, they are limited in their ability to reveal how these words are learned, which calls for experiments to

test causal hypotheses. To complement this, Chapters 4 and 5 reported word learning experiments to test how words are learned in emotional context; and while these two chapters focused more on the role of the narrative context, emotionality remained a key aspect, particularly in terms of the overall positivity or negativity of language. The narrative stimuli were evaluated for their sentiments using *BERT* similar to that employed in Chapter 3, and responses from the meaning generation task were assessed using the valence norms established by Warriner et al. (2013). These experimental chapters demonstrate that adults and children were able to learn new adjectives from reading, and that emotional contexts better support the learning of new words compared to neutral contexts. Finally, Chapter 6 integrates different approaches to conceptualising emotions, examining children's knowledge and use of emotion words in isolated production as well as in narrative contexts. This chapter also explores the general emotional tone of language in narrative production through sentiment analysis. Engaging with multiple approaches to understanding emotion language captures its inherent complexity. It also highlights the value of exploring this phenomenon from various perspectives.

Additionally, as set out in the thesis scope in Chapter 1, the broader level consideration of emotion words focused on the dimension of valence. While valence is a critical aspect of emotional experience, and one that develops the earliest (Widen & Russell, 2008), other affective and lexical features such as arousal, dominance, and concreteness might also play important roles in shaping emotional experiences and word learning. For example, in the experimental stimuli in Chapters 4 and 5, arousal and concreteness were not strictly controlled. This was a calculated trade-off aimed at balancing the use of naturalistic language stimuli with the need for experimental control. While strictly controlling for multiple emotional dimensions could have enhanced precision, it might have also introduced artificiality as arousal is inherently a characteristic of emotionality. Yet, at the same time, I recognise that this methodological choice presents challenges in teasing apart the influence of arousal in explaining the better performance in emotional contexts as compared to neutral. This trade-off highlights a broader reflection on designing psycholinguistic experiments: how to connect results from tightly controlled experiments that isolate specific variables of interest to the complex and

natural reading processes that people experience in real life.

Emotion language, emotion understanding, and mental health

The constructionist approach to emotional development, reviewed in Chapter 1 and throughout the introductory sections of each chapter, highlights the critical role of emotion language and particularly emotion words in the acquisition of emotional concepts and understanding. This connection provides a key motivation for my investigation of emotion language. My thesis offers multiple perspectives on the frequency and diversity of emotion language in children's language environments and their own production.

While not directly tested in my thesis, emotion language has been proposed to hold potential as both a diagnostic tool and a target for mental health intervention (Nook, 2023). As a diagnostic tool, sentiment analysis methods, like those used in Chapters 3 to 6, assess the emotional valence of language, which could offer a proxy for an individual's emotional well-being (Pennebaker et al., 2003). These measures of sentiments have been reported to correlate with self-reported affect and to predict mental health symptoms such as depression and anxiety (e.g., Du & Sun, 2022; Eberhardt et al., 2024; Silveira et al., 2021). Yet, the effectiveness often depends on the model used, and inconsistencies in the literature as well as considerations regarding transparency and ethics highlight the need for more robust cross-validation (e.g., Hartmann et al., 2023; T. Zhang et al., 2022). Beyond diagnosis, emotion language could also be used for intervention. Techniques like affect labelling, which is the process of putting one's feelings into words, have been shown to implicitly downregulate emotional responses (e.g., Constantinou et al., 2014; Kircanski et al., 2012; Torre & Lieberman, 2018). Additionally, linguistic distancing, which involves shifting language from first-person ("I am upset") to third-person ("She is upset"), creates psychological distance from the emotional mental states and has been linked to reduction in the intensity of negative emotions and improved mental health outcomes (e.g., Kross et al., 2014). However, not all studies reported positive results in adults, and very few studies examined their effectiveness in children (e.g., Marks et al., 2019; Plaisted et al., 2022).

Admittedly, children's developing language ability and often immature understanding of emotion concepts adds complexity. For example, a child might experience frustration but lack the linguistic tools to articulate it, leading to behavioural expressions like tantrums. Conversely, a child might verbally label an emotion like "sadness" yet struggle to express or cope with it appropriately. This classic distinction between knowledge and performance is crucial, as it highlights that lexical knowledge and emotional comprehension do not always develop in tandem. The effectiveness of language as a crucial tool for emotion understanding is influenced by numerous other factors associated with the changing social environment and developing brain (Nook, 2023). Emotional vocabulary alone may not be sufficient to foster mental well-being unless it is integrated with broader social communication competencies (Rimehaug & Kårstad, 2022).

Reflecting on my thesis, while the importance of emotion language is evident, I did not directly assess children's emotions by using measures of emotion recognition or emotion regulation, for example. The three pieces to the puzzle, knowing an emotion, having the words for it, and effectively using the words to support emotional functioning and mental health, are each important in its own right; and the relationship between the use of emotion language and emotional understanding remains an open question. This highlights the need for further research to investigate the nuances of emotion language and how these learned, understood, and used by children of different ages. Understanding these associations more deeply is an important and exciting research agenda, and one that I would love to explore in the future. Such research could ultimately provide better insights into the interplay between language and emotion in human communication and development.

Additionally, despite not a direct focus of my thesis, it is important to recognise and support children with varying language abilities, especially those who struggle with language. These children may lag behind in acquiring emotion language through books, which can, in turn, affect their social-emotional development and mental health (e.g., S. Griffiths et al., 2024; Jelen et al., 2023; Wilmot et al., 2023). Interventions such as interactive storytelling or multimodal resources could potentially make emotion language more accessible, fostering emotion understanding and resilience in a diverse range of

learners. These insights hold practical significance for improving social-emotional skills and well-being, and ultimately quality of life across diverse populations.

Rethinking contexts and generalisability

The importance of context in emotion language is emphasised throughout this thesis and reflected in its title. Context in my thesis can be understood in three dimensions.

Contextualised in narratives

The first dimension involves understanding a word in the context of its surrounding words to grasp its meaning. Chapters 3 to 6 all examined emotion language within narrative context. The importance of considering language context beyond isolated words has been discussed in the earlier section of this chapter.

The insight of corpus analysis often depends on the availability and quality of the corpus. The computation and analyses of children's written narratives in Chapter 3 (sentiment analysis) and in Chapter 5 (context valence) was made possible by the Oxford University Press kindly providing access to the children's writing corpus. This has allowed me to explore the contexts in which children use emotion words. However, due to copyright restrictions, it was more challenging to access the full text of the children's reading corpus and the television subtitle corpus, for which I could only work with frequency lists. Thus, this limitation reduced the analysis in Chapter 2 to a "bag of words" approach, which stripped the words of their contextual richness and constrained the insights I could draw regarding the use of emotion words in context. Nevertheless, documenting and comparing the frequency and diversity of these emotion words remain valuable and meaningful, as it provides the foundation for exploring other important questions in future research.

I also reflected on the balance between using prompts and allowing free narratives when collecting children's narrative production. In Chapter 3, one of the advantages of the children's writing corpus was that there were no constraints on topics, which allowed me to analyse children's free writing and

to explore the topics they naturally wanted to write about. I then aimed to mirror this approach in Chapter 6 with spoken language production, where children were free to talk about any memorable day. Yet, as discussed in Chapter 6, this approach might have complicated analyses and limited comparability across participants and contexts. The nature of spoken language introduces greater variability, and the diversity of topics, especially given the relatively small sample size (compared to the corpus), makes it harder to detect meaningful trend in the use of emotion words. This variability raises the question of whether complete freedom in language production, while offering insights into natural language use and preference, is optimal for the purpose of psycholinguistic research. It is a balance that future research should carefully consider.

Contextualised in written language

The second dimension is the context of written language, rather than conversational language, and how this becomes increasingly important as children's literacy kicks in. Mid-childhood marks an important period for learning through reading (Nation et al., 2022). The significance of written language is demonstrated across several chapters in this thesis. For instance, the cross-corpus comparison in Chapter 2 shows that written language targeted at children contains more words related to emotions and mental states than age-matched child-directed speech. The experiment reported in Chapter 5 found that children can infer the emotional associations of novel words from short narratives, highlighting reading as a key avenue for language learning. Furthermore, Chapter 6 reports a positive correlation between children's exposure to book language and the lexical diversity of their narrative production. Although these correlational findings from corpus analyses and cross-sectional samples could not imply causation, it is likely that the language children experience through reading provides a rich substrate from which children can learn emotion vocabulary and word valence.

While this thesis emphasises the greater lexical diversity of book language, it does not dismiss the importance of spoken interactions in contributing to the development of emotion language. Spoken language, and especially shared conversational language, is enriched by prosody, intonation, and the shared situational context between speaker and listener. When interaction is present, these

conversational exchanges in child-directed speech might also contribute to children's emotional understanding by offering immediate, contextually grounded examples. These spoken interactions and multimodal engagement might be especially important for children with language difficulties.

In contrast, written language is crafted to stand independently, requiring the reader to reconstruct the intended situation using only the textual information provided. Particularly in books, the careful selection of words by the author enables a richer lexical diversity, often incorporating advanced emotion words not commonly found in everyday speech. For example, Chapter 2 shows that words like "appalled" appear in picture books but not in child-directed speech or television programmes. The ability of written language to convey diverse emotional scenarios through text alone makes it a valuable medium for expanding a child's emotional vocabulary.

It is worth noting that there are various forms of written language. For example, academic language, storybook language and nonfictions may differ in their word choice and sentence structure. Narrative fictions have been reported to have a larger proportion of complex emotions (e.g., amusement, despair) as well as words used in an emotive sense (e.g., a sigh of relief vs. hurricane relief fund) than in nonfiction texts (Schwering et al., 2021). Exposure to narrative fictions were found to be a better predictor for verbal abilities and emotional development than expository nonfiction (e.g., Castano, 2024; Mar & Rain, 2015). Castano (2024) proposed the less-is-more hypothesis, which suggested that emotion recognition skills are strengthened by stories that do not *tell* us, but rather *show* us the emotional life of fictional characters. My thesis focuses mainly on narrative fictions written for and by children, yet these different forms of written language might offer different learning opportunities for children, which are beyond the scope of the thesis.

Future research could explore how different forms of written and spoken language may complement each other in the development of emotion language. It would also be interesting to investigate whether the emotional nuance and immediacy of spoken interactions may enhance the comprehension of more complex emotion words encountered in written language. On this view, shared book reading is particularly valuable for developing language skills for preschool children before they can read

independently (Fitton et al., 2018; Noble et al., 2020), as it incorporates both greater lexical diversity of written narratives and the interactive element in conversations, offering the opportunity to discuss meaning in context. Additionally, shared reading introduces children to literacy, getting them interested in reading and motivated to learn how to read (Seidenberg & MacDonald, 2018). As children's reading skills mature, independent reading then offers greater efficiency to access a broader range of text in a shorter time, further supporting children's emotion and language development.

Contextualised in English

The third dimension of context pertains to the language itself, that my thesis is contextualised in English, and more precisely, British English. This focus does not capture the full spectrum of emotional nuances present in other languages. English dominates psychological and cognitive research in language and emotions, which offers readily available lexicons of word-affect associations and machine learning algorithms for analysis. However, this dominance also limits the generalisability of findings. For example, the majority of the training data used in large language models (LLMs) is English, which inherently biases them toward English norms and values. This is a limitation that permeates much of the existing language research. Nevertheless, data-driven approaches to analysing language corpora can reveal patterns in how people use language, and from this, help address psychological questions (Jackson et al., 2022, p. 807). English should serve as a starting point for language research but not the boundary.

There are important cultural variations in how emotions are categorised and expressed, which might pose significant challenges to generalising findings about emotions (Majid, 2022; Wierzbicka, 2009). For example, Lutz (1982) found that on the island of Ifaluk, emotion words are categorised based on social situations, reflecting the cultural context. Even among closely related cultures, such as the six Western European countries studied, Van Goozen and Frijda (1993) reported considerable differences in the relative frequency of presumably equivalent emotion words, despite broad similarities in emotion categories. Even basic emotions that are often proposed as universally experienced can be articulated and interpreted differently across cultures. For instance, some societies express emotions

like sadness or grief through stoic behaviour, reflecting cultural norms that value restraint. In contrast, other cultures may encourage more outward displays of these emotions, leading to differing interpretations of emotional states. Ng et al. (2019) explored the organisation and distribution of emotion words in Mandarin Chinese, revealing cross-linguistic divergences in valence and intensity that challenge the universality of emotion word associations. Furthermore, the quest for psychological universals is often impeded by methodological limitations, particularly the overestimation of cultural diversity within study samples, as critiqued by Majid (2023). The exploration of emotion and language thus reveals a complex interplay between universal psychological mechanisms and culturally specific expressions. Addressing these complexities through rigorous cross-cultural and cross-linguistic research will lead to a more nuanced understanding of the universality and diversity inherent in human emotion and language (e.g., Bergelson et al., 2023; Jackson et al., 2019, 2022).

Limitations of the thesis

Having reflected on each key theme of the thesis, I now review and summarise its limitations. First, the thesis exclusively focused on emotion language as expressed in text or transcribed speech. It did not involve any direct measurement of children's emotional experiences or their comprehension of emotions. As a result, this thesis is unable to establish a direct link between the usage of emotion words and actual emotional expression or mental health outcomes in children. Future research should aim to bridge this gap by incorporating assessments of emotional experience and comprehension alongside language measures.

When examining emotion words in children's language corpora, the thesis focused on a limited subset of emotion words, potentially overlooking many other emotion words that children use. Additionally, emotions are conveyed not only through words but also through prosody, metaphors, grammar, and punctuation, as suggested by Majid (2012). Future studies should integrate these diverse linguistic features to capture a more comprehensive picture of emotional expression in language. Moreover, the corpora used may not fully represent British children's experiences. For example, the writing corpus was built from national competitions, likely children with stronger language skills, potentially

excluding populations with weaker language abilities. This limitation highlights the need for broader sampling strategies to better reflect variability in children's language development and emotional expression.

In the experimental sections, the thesis primarily interpreted findings through the affective embodiment account. However, alternative mechanisms, such as distributed semantics and statistical learning, might also partly explain the results but were not explored in depth. Furthermore, word learning was conceptualised in a relatively simplified manner, focusing on the acquisition of word form and valence. Whether this constitutes comprehensive word learning is open to debate. A clearer definition of "word learning" at the outset would improve interpretability and facilitate comparisons with other existing studies. Moreover, the meaning of novel words was not provided in the experiments, raising uncertainties about whether children could retain the meanings of these words and generalise their usage to new contexts. Future research should assess word knowledge after a delay to determine whether learning effects persist over time and employ additional tasks to assess word meaning learning beyond valence. Another limitation concerns the methodological choice not to control for arousal. This presents challenges in disentangling the specific influence of arousal from valence in explaining children's improved performance in emotional contexts compared to neutral ones. Addressing this limitation in future studies could provide a more nuanced understanding of the mechanisms underlying these effects.

Finally, while the thesis emphasises the role of written language and reading in children's language development, it recognises that reading is not the sole source of input for learning emotion words. Children acquire emotion language from a variety of sources, including conversations with peers, school curricula, and other interactions beyond reading. Future research should adopt a more holistic approach to examine how these diverse sources collectively contribute to children's language and emotional development.

Concluding remarks

Emotions are closely associated with language in both adults and children. My thesis examined emotion language in children's reading, writing, and word learning from multiple perspectives. Data-driven corpus analyses studied the frequency and distribution of emotion words in texts both written by and for children, as well as the emotional valence associated with these words. My results demonstrated that language written by and for older children contained a greater diversity of emotion words than spoken language targeting at younger children. My word learning experiments investigated whether the emotional context within narratives influences word learning through reading. Both adults and children were able to infer the valence of novel words from reading short narratives, and those words read in emotional contexts were learned better. In sum, by combining data-driven analyses and experimental work, this thesis offers a deeper understanding of the emotional content of language written by and for children, and how emotional contexts influence word learning. The theoretical evidence provided by the thesis holds implications for emotional well-being, language and literacy development, and educational practice.

References

Abeare, C. A., An, K., Tyson, B., Holcomb, M., Cutler, L., May, N., & Erdodi, L. A. (2022). The emotion word fluency test as an embedded performance validity indicator – Alone and in a multivariate validity composite. *Applied Neuropsychology: Child*, *11*(4), 713–724. https://doi.org/10.1080/21622965.2021.1939027

Abeare, C. A., Freund, S., Kaploun, K., McAuley, T., & Dumitrescu, C. (2017). The Emotion Word Fluency Test (EWFT): Initial psychometric, validation, and physiological evidence in young adults. *Journal of Clinical and Experimental Neuropsychology*, *39*(8), 738–752. https://doi.org/10.1080/13803395.2016.1259396

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289. https://doi.org/10.3758/BRM.40.1.278

Adrian, J. E., Clemente, R. A., Villanueva, L., & Rieffe, C. (2005). Parent–child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, *32*(3), 673–686. https://doi.org/10.1017/S0305000905006963

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Bahn, D., Vesker, M., García Alanis, J. C., Schwarzer, G., & Kauschke, C. (2017). Age-Dependent Positivity-Bias in Children's Processing of Emotion Terms. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01268

Bailen, N. H., Green, L. M., & Thompson, R. J. (2019). Understanding Emotion in Adolescents: A Review of Emotional Frequency, Intensity, Instability, and Clarity. *Emotion Review*, *11*(1), 63–73. https://doi.org/10.1177/1754073918768878

Baron-Cohen, S., Golan, O., Wheelwright, S., & Granader, Y. (2010). Emotion Word Comprehension from 4 to 16 Years Old: A Developmental Survey. *Frontiers in Evolutionary Neuroscience*, *2*. https://www.frontiersin.org/articles/10.3389/fnevo.2010.00109

Barrett, L. F. (2004). Feelings or Words? Understanding the Content in Self-Report Ratings of Experienced Emotion. *Journal of Personality and Social Psychology*, *87*(2), 266–281. https://doi.org/10.1037/0022-3514.87.2.266

Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. https://doi.org/10.1093/scan/nsw154

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Batini, F., Luperini, V., Cei, E., Izzo, D., & Toti, G. (2021). The Association Between Reading and Emotional Development: A Systematic Review. *Journal of Education and Training Studies*, *9*(1), Article 1. https://doi.org/10.11114/jets.v9i1.5053

Bayer, M., Sommer, W., & Schacht, A. (2010). Reading emotional words within sentences: The impact of arousal and valence on event-related potentials. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *78*(3), 299–307. https://doi.org/10.1016/j.ijpsycho.2010.09.004

Beck, L., Kumschick, I. R., Eid, M., & Klann-Delius, G. (2012). Relationship between language competence and emotional competence in middle childhood. *Emotion (Washington, D.C.)*, *12*(3), 503–514. https://doi.org/10.1037/a0026320

Bergelson, E., Soderstrom, M., Schwarz, I.-C., Rowland, C. F., Ramírez-Esparza, N., R. Hamrick, L., Marklund, E., Kalashnikova, M., Guez, A., Casillas, M., Benetti, L., Alphen, P. van, & Cristia, A. (2023). Everyday language input and production in 1,001 children from six continents. *Proceedings of the National Academy of Sciences*, *120*(52), e2300671120. https://doi.org/10.1073/pnas.2300671120

Bireta, T. J., Guitard, D., Neath, I., & Surprenant, A. M. (2023). Valence and concreteness in item recognition: Evidence against the affective embodiment account. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02442-8

Bishop, D. V. M. (2004). *ERRNI: Expression, Reception and Recall of Narrative Instrument*. Harcourt Assessment.

Boerma, I. E., Mol, S. E., & Jolles, J. (2017). The Role of Home Literacy Environment, Mentalizing, Expressive Verbal Ability, and Print Exposure in Third and Fourth Graders' Reading Comprehension. *Scientific Studies of Reading*, *21*(3), 179–193. https://doi.org/10.1080/10888438.2016.1277727

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*(2), 122–159. https://doi.org/10.1080/01638530701792826

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, *143*(3), 263–292. https://doi.org/10.1037/bul0000089

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning & Verbal Behavior*, *8*, 1–8. https://doi.org/10.1016/S0022-5371(69)80002-2

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin & Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Boyes, M. E., Leitao, S., Claessen, M., Badcock, N. A., & Nayton, M. (2016). Why Are Reading Difficulties Associated with Mental Health Problems? *Dyslexia*, *22*(3), 263–266. https://doi.org/10.1002/dys.1531

Boyes, M. E., Tebbutt, B., Preece, K. A., & Badcock, N. A. (2018). Relationships between Reading Ability and Child Mental Health: Moderating Effects of Self-Esteem. *Australian Psychologist*, *53*(2), 125–133. https://doi.org/10.1111/ap.12281

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*(2), 218–224. https://doi.org/10.1016/j.lindif.2010.10.002

Brierley, B., Medford, N., Shaw, P., & David, A. S. (2007). Emotional memory for words: Separating content and context. *Cognition and Emotion*, *21*(3), 495–521. https://doi.org/10.1080/02699930600684963

Brysbaert, M., Sui, L., Dirix, N., & Hintz, F. (2020). Dutch Author Recognition Test. *Journal of Cognition*, *3*(1). https://doi.org/10.5334/joc.95

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Camodeca, A., Walcott, K., Hosack, A., & Todd, K. Q. (2021). Preliminary evidence for the Emotion Word Fluency Test as a unique semantic fluency measure. *Psychological Assessment*, *33*(2), 195–200. https://doi.org/10.1037/pas0000965

Castano, E. (2024). Less Is More: How the Language of Fiction Fosters Emotion Recognition. *Emotion Review*, *16*(2), 73–83. https://doi.org/10.1177/17540739241232350

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*(4), 735–765. https://doi.org/10.1037/a0030737

Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender Stereotypes in Natural Language: Word Embeddings Show Robust Consistency Across Child and Adult Language Corpora of More Than 65 Million Words. *Psychological Science*, *32*(2), 218–240. https://doi.org/10.1177/0956797620963619

Chilton, M. W., & Ehri, L. C. (2015). Vocabulary Learning: Sentence Contexts Linked by Events in Scenarios Facilitate Third Graders' Memory for Verb Meanings. *Reading Research Quarterly*, *50*(4), 439–458. https://doi.org/10.1002/rrq.106

Cohen, K. A., Shroff, A., Nook, E. C., & Schleider, J. L. (2022). Linguistic distancing predicts response to a digital single-session intervention for adolescent depression. *Behaviour Research and Therapy*, *159*, 104220. https://doi.org/10.1016/j.brat.2022.104220

Constantinou, E., Van Den Houte, M., Bogaerts, K., Van Diest, I., & Van den Bergh, O. (2014). Can words heal? Using affect labeling to reduce the effects of unpleasant cues on symptom reporting. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00807

Crossfield, E., & Damian, M. F. (2021). The role of valence in word processing: Evidence from lexical decision and emotional Stroop tasks. *Acta Psychologica*, *218*, 103359. https://doi.org/10.1016/j.actpsy.2021.103359

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Cunningham, W. A., Dunfield, K. A., & Stillman, P. E. (2013). Emotional states from affective dynamics. *Emotion Review*, *5*(4), 344–355. https://doi.org/10.1177/1754073913489749

Davies, C., Lingwood, J., & Arunachalam, S. (2020). Adjective forms and functions in British English child-directed speech. *Journal of Child Language*, *47*(1), 159–185. https://doi.org/10.1017/S0305000919000242

Davies, C., Syrett, K., Taylor, L., Wilkes, S., & Zuniga-Montanez, C. (2022). Supporting adjective learning across the curriculum by 5–7 year-olds: Insights from psychological research. *Language and Linguistics Compass*, *16*(11), e12476. https://doi.org/10.1111/lnc3.12476

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Dawson, N., Hsiao, Y., Tan, A. W. M., Banerji, N., & Nation, K. (2021). Features of lexical richness in children's books: Comparisons with child-directed speech. *Language Development Research*, *1*(1), Article 1. https://doi.org/10.34842/5we1-yk94

Dawson, N., Hsiao, Y., Tan, A. W. M., Banerji, N., & Nation, K. (2023). Effects of Target Age and Genre on Morphological Complexity in Children's Reading Material. *Scientific Studies of Reading*, *27*(6), 529–556. https://doi.org/10.1080/10888438.2023.2206574

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Dockrell, J. E., & Connelly, V. (2016). The Relationships Between Oral and Written Sentence Generation in English Speaking Children: The Role of Language and Literacy Skills. In J. Perera, M. Aparici, E. Rosado, & N. Salas (Eds.), *Written and Spoken Language Development across the Lifespan: Essays in Honour of Liliana Tolchinsky* (pp. 161–177). Springer International Publishing. https://doi.org/10.1007/978-3-319-21136-7_11

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdoomian, K., McMahon, M. T., Tivnan, B. F., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, *112*(8), 2389–2394. https://doi.org/10.1073/pnas.1411678112

Dong, Y., Hsiao, Y., Dawson, N., Banerji, N., & Nation, K. (2024). The Emotional Content of Children's Writing: A Data-Driven Approach. *Cognitive Science*, *48*(3), e13423. https://doi.org/10.1111/cogs.13423

Dong, Y., Mak, M. H., Hepach, R., & Nation, K. (2024). Learning new words via reading: The influence of emotional narrative context on learning novel adjectives. *Quarterly Journal of Experimental Psychology*, *0*(0). https://doi.org/10.1177/17470218241308221

Dong, Y., & Nation, K. (2024). *Charting the frequency and diversity of emotion words in children's language: Written language matters*. OSF. https://doi.org/10.31219/osf.io/zj5k7

Doost, H. T. N., Moradi, A. R., Taghavi, M. R., Yule, W., & Dalgleish, T. (1999). The development of a corpus of emotional words produced by children and adolescents. *Personality and Individual Differences*, *27*, 433–451. https://doi.org/10.1016/S0191-8869(98)00253-0

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. *Journal of Experimental Psychology: General*, *147*(1), 62–73. https://doi.org/10.1037/xge0000361

Driver, M. (2021). EMOTION-LADEN TEXTS AND WORDS: THE INFLUENCE OF EMOTION ON VOCABULARY LEARNING FOR HERITAGE AND FOREIGN LANGUAGE LEARNERS. *Studies in Second Language Acquisition*, 1–24. https://doi.org/10.1017/S0272263121000851

Du, X., & Sun, Y. (2022). Linguistic features and psychological states: A machine-learning based approach. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.955850

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology*, *27*(3), 448–455. https://doi.org/10.1037/0012-1649.27.3.448

Eberhardt, S. T., Schaffrath, J., Moggia, D., Schwartz, B., Jaehde, M., Rubel, J. A., Baur, T., André, E., & Lutz, W. (2024). Decoding emotions: Exploring the validity of sentiment analysis in psychotherapy. *Psychotherapy Research*, *0*(0), 1–16. https://doi.org/10.1080/10503307.2024.2322522

Egbert, J., Gray, B., & Biber, D. (Eds.). (2022). Distribution Considerations. In *Designing and Evaluating Language Corpora: A Practical Framework for Corpus Representativeness* (pp. 122–155). Cambridge University Press. https://doi.org/10.1017/9781316584880.005

Eggum, N. D., Eisenberg, N., Kao, K., Spinrad, T. L., Bolnick, R., Hofer, C., Kupfer, A. S., & Fabricius, W. V. (2011). Emotion understanding, theory of mind, and prosocial orientation: Relations over time in early childhood. *The Journal of Positive Psychology*, *6*(1), 4–16. https://doi.org/10.1080/17439760.2010.536776

Eke, C. I., Norman, A. A., & Shuib, L. (2021). Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model. *IEEE Access*, *9*, 48501–48518. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3068323

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition*, *108*(2), 557–565. https://doi.org/10.1016/j.cognition.2008.03.003

Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric Evaluation of Lexical Diversity Indices: Assessing Length Effects. *Journal of Speech, Language, and Hearing Research : JSLHR*, *58*(3), 840–852. https://doi.org/10.1044/2015_JSLHR-L-14-0280

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889–906. https://doi.org/10.1037/0022-3514.38.6.889

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared Book Reading Interventions With English Learners: A Meta-Analysis. *Review of Educational Research*, *88*(5), 712–751. https://doi.org/10.3102/0034654318790909

Fong, K., Mullin, J., & Mar, R. (2013). What You Read Matters: The Role of Fiction Genre in Predicting Interpersonal Sensitivity. *Psychology of Aesthetics Creativity and the Arts*, *7*. https://doi.org/10.1037/a0034084

Frances, C., Bruin, A. de, & Duñabeitia, J. A. (2020). The effects of language and emotionality of stimuli on vocabulary learning. *PLOS ONE*, *15*(10), e0240252. https://doi.org/10.1371/journal.pone.0240252

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. https://doi.org/10.1017/S0305000916000209

Frost, A., Hoyt, L. T., Chung, A. L., & Adam, E. K. (2015). Daily life with depressive symptoms: Gender differences in adolescents' everyday emotional experiences. *Journal of Adolescence*, *43*, 132–141. https://doi.org/10.1016/j.adolescence.2015.06.001

Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proceedings of the National Academy of Sciences*, *114*(30), 7900–7907. https://doi.org/10.1073/pnas.1621073114

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*(2), 314–325. https://doi.org/10.1037/a0026007

Gowenlock, A., Norbury, C., & Rodd, J. M. (2024). *Exposure to Language in Video and its Impact on Linguistic Development in Children Aged 3-11: A Scoping Review*. https://osf.io/5e24g

Graham, S. (2006). Writing. In *Handbook of educational psychology* (pp. 457–478). Lawrence Erlbaum Associates Publishers.

Graham, S. (2018). Introduction to Conceptualizing Writing. *Educational Psychologist*, *53*(4), 217–219. https://doi.org/10.1080/00461520.2018.1514303

Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, *27*(2), 99–111. https://doi.org/10.1016/j.cogdev.2012.03.002

Green, C., Keogh, K., Sun, H., & O'Brien, B. (2023). The Children's Picture Books Lexicon (CPB-Lex): A large-scale lexical database from children's picture books. *Behavior Research Methods*. https://doi.org/10.3758/s13428-023-02198-y

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral Discourse in the Preschool Years and Later Literacy Skills. *First Language*, *24*(2), 123–147. https://doi.org/10.1177/0142723704042369

Griffiths, S., Goh, S. K. Y., & Norbury, C. F. (2020). Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues. *PeerJ*, *8*, e9118. https://doi.org/10.7717/peerj.9118

Griffiths, S., Lucas, L., Gooch, D., & Norbury, C. F. (2024). Special educational needs provision and academic outcomes for children with teacher reported language difficulties at school entry. *JCPP Advances*, *n/a*(n/a), e12264. https://doi.org/10.1002/jcv2.12264

Griffiths, S., Suksasilp, C., Lucas, L., Sebastian, C. L., Norbury, C., & the SCALES team. (2021). Relationship between early language competence and cognitive emotion regulation in adolescence. *Royal Society Open Science*, *8*(10), 210742. https://doi.org/10.1098/rsos.210742

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*, 211–244. https://doi.org/10.1037/0033-295X.114.2.211

Grosse, G., Streubel, B., Gunzenhauser, C., & Saalbach, H. (2021). Let's Talk About Emotions: The Development of Children's Emotion Vocabulary from 4 to 11 Years of Age. *Affective Science*, *2*(2), 150–162. https://doi.org/10.1007/s42761-021-00040-2

Hagelskamp, C., Brackett, M. A., Rivers, S. E., & Salovey, P. (2013). Improving Classroom Quality with The RULER Approach to Social and Emotional Learning: Proximal and Distal Outcomes. *American Journal of Community Psychology*, *51*(3–4), 530–543. https://doi.org/10.1007/s10464-013-9570-x

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In *Handbook of metacognition in education* (pp. 131–153). Routledge/Taylor & Francis Group.

Hartmann, J., Heitmann, M., Siebert, C., & Schamp, C. (2023). More than a Feeling: Accuracy and Application of Sentiment Analysis. *International Journal of Research in Marketing*, *40*(1), 75–87. https://doi.org/10.1016/j.ijresmar.2022.05.005

Hastie, T., & Tibshirani, R. (1990). Exploring the Nature of Covariate Effects in the Proportional Hazards Model. *Biometrics*, *46*(4), 1005–1016. https://doi.org/10.2307/2532444

Hegefeld, H. M., Satpute, A. B., Ochsner, K. N., Davidow, J. Y., & Nook, E. C. (2023). Fluency generating emotion words correlates with verbal measures but not emotion regulation, alexithymia, or depressive symptoms. *Emotion (Washington, D.C.)*, *23*(8), 2259–2269. https://doi.org/10.1037/emo0001229

Hipson, W., & Mohammad, S. M. (2020). PoKi: A Large Dataset of Poems by Children. *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 1578–1589. https://aclanthology.org/2020.lrec-1.196

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, *55*(9), 1830–1849. https://doi.org/10.1037/dev0000686

Hogan, P. C. (2011). *What Literature Teaches Us about Emotion*. Cambridge University Press. https://doi.org/10.1017/CBO9780511976773

Hsiao, Y., Banerji, N., & Nation, K. (2021). Boys Write About Boys: Androcentrism in Children's Reading Experience and Its Emergence in Children's Own Writing. *Child Development*, *92*(6), 2194–2204. https://doi.org/10.1111/cdev.13623

Hsiao, Y., Bird, M., Norris, H., Pagán, A., & Nation, K. (2020). The influence of item-level contextual history on lexical and semantic judgments by children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2367–2383. https://doi.org/10.1037/xlm0000795

Hsiao, Y., Dawson, N. J., Banerji, N., & Nation, K. (2023). The nature and frequency of relative clauses in the language children hear and the language children read: A developmental cross-corpus analysis of English complex grammar. *Journal of Child Language*, *50*(3), 555–580. https://doi.org/10.1017/S0305000921000957

Hsiao, Y., Dawson, N. J., Banerji, N., & Nation, K. (2024). A corpus-based developmental investigation of linguistic complexity in children's writing. *Applied Corpus Linguistics*, *4*(1), 100084. https://doi.org/10.1016/j.acorp.2024.100084

Hulme, R. C., Begum, A., Nation, K., & Rodd, J. M. (2023). Diversity of narrative context disrupts the early stage of learning the meanings of novel words. *Psychonomic Bulletin & Review*, *30*(6), 2338–2350. https://doi.org/10.3758/s13423-023-02316-z

Jackson, J. C., Lindquist, K., Drabble, R., Atkinson, Q., & Watts, J. (2023). Valence-dependent mutation in lexical evolution. *Nature Human Behaviour*, *7*(2), Article 2. https://doi.org/10.1038/s41562-022-01483-8

Jackson, J. C., Watts, J., Henry, T. R., List, J.-M., Forkel, R., Mucha, P. J., Greenhill, S. J., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science*, *366*(6472), 1517–1522. https://doi.org/10.1126/science.aaw8160

Jackson, J. C., Watts, J., List, J.-M., Puryear, C., Drabble, R., & Lindquist, K. A. (2022). From Text to Thought: How Analyzing Language Can Advance Psychological Science. *Perspectives on Psychological Science*, *17*(3), 805–826. https://doi.org/10.1177/17456916211004899

Jelen, M. B., Griffiths, S. L., Lucas, L., Saul, J., & Norbury, C. F. (2023). The role of language in mental health during the transition from primary to secondary education. *Quarterly Journal of Experimental Psychology*, *76*(12), 2732–2748. https://doi.org/10.1177/17470218231158069

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, *166*, 190–211. https://doi.org/10.1016/j.jecp.2017.08.010

Joseph, H. S. S. L., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, *133*(1), 238–248. https://doi.org/10.1016/j.cognition.2014.06.015

Kanaris, A. (1999). Gendered Journeys: Children's Writing and the Construction of Gender. *Language and Education*, *13*(4), 254–268. https://doi.org/10.1080/09500789908666772

Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). Unpacking Emotion Differentiation: Transforming Unpleasant Experience by Perceiving Distinctions in Negativity. *Current Directions in Psychological Science*, 7. https://doi.org/10.1177/0963721414550708

Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The Role of Emotional Valence for the Processing of Facial and Verbal Stimuli—Positivity or Negativity Bias? *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01654

Kidd, D., & Castano, D. (2013). Reading Literary Fiction Improves Theory of Mind. *Science*, *342*(6156), 377–380. https://doi.org/10.1126/science.1239918

Kidd, D., & Castano, E. (2019). Reading Literary Fiction and Theory of Mind: Three Preregistered Replications and Extensions of Kidd and Castano (2013). *Social Psychological and Personality Science*, *10*(4), 522–531. https://doi.org/10.1177/1948550618775410

Kim, J. M., Sidhu, D. M., & Pexman, P. M. (2020). Effects of Emotional Valence and Concreteness on Children's Recognition Memory. *Frontiers in Psychology*, *11*, 615041. https://doi.org/10.3389/fpsyg.2020.615041

Kim, Y.-S. G., Park, C., & Park, Y. (2015). Dimensions of discourse level oral language skills and their relation to reading comprehension and written composition: An exploratory study. *Reading and Writing*, *28*(5), 633–654. https://doi.org/10.1007/s11145-015-9542-7

Kircanski, K., Lieberman, M. D., & Craske, M. G. (2012). Feelings into words: Contributions of language to exposure therapy. *Psychological Science*, *23*(10), 1086–1091. https://doi.org/10.1177/0956797612443830

Kolak, J., Monaghan, P., & Taylor, G. (2023). Language in educational apps for pre-schoolers. A comparison of grammatical constructions and psycholinguistic features in apps, books and child directed speech. *Journal of Child Language*, *50*(4), 895–921. https://doi.org/10.1017/S0305000922000198

Korochkina, M., Marelli, M., Brysbaert, M., & Rastle, K. (2024). The Children and Young People's Books Lexicon (CYP-LEX): A large-scale lexical database of books read by children and young people in the United Kingdom. *Quarterly Journal of Experimental Psychology*, 17470218241229694. https://doi.org/10.1177/17470218241229694

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*(1), 14–34. https://doi.org/10.1037/a0021446

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Kross, E., Bruehlman-Senecal, E., Park, J., Burson, A., Dougherty, A., Shablack, H., Bremner, R., Moser, J., & Ayduk, O. (2014). Self-talk as a regulatory mechanism: How you do it matters. *Journal of Personality and Social Psychology*, *106*(2), 304–324. https://doi.org/10.1037/a0035173

Kucirkova, N. (2019). How Could Children's Storybooks Promote Empathy? A Conceptual Framework Based on Developmental Psychology and Literary Theory. *Frontiers in Psychology*, *10*, 121. https://doi.org/10.3389/fpsyg.2019.00121

Kuhlmann, M., Hofmann, M. J., & Jacobs, A. M. (2017). If You Don't Have Valence, Ask Your Neighbor: Evaluation of Neutral Words as a Function of Affective Semantic Associates. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00343

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology. General*, *143*(3), 1065–1081. https://doi.org/10.1037/a0035669

Lana, N., & Kuperman, V. (2023). Learning concrete and abstract novel words in emotional contexts: Evidence from incidental vocabulary learning. *Language Learning and Development*, *20*(2), 158-173. https://doi.org/10.1080/15475441.2023.2246438

Larson, R. W., Moneta, G., Richards, M. H., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*, *73*(4), 1151–1165. https://doi.org/10.1111/1467-8624.00464

Lenci, A., Lebani, G. E., & Passaro, L. C. (2018). The Emotions of Abstract Words: A Distributional Semantic Analysis. *Topics in Cognitive Science*, *10*(3), 550–572. https://doi.org/10.1111/tops.12335

Lewis, M., Cooper Borkenhagen, M., Converse, E., Lupyan, G., & Seidenberg, M. S. (2022). What Might Books Be Teaching Young Children About Gender? *Psychological Science*, *33*(1), 33–47. https://doi.org/10.1177/09567976211024643

Lewis, M., & Lupyan, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. *Nature Human Behaviour*, *4*(10), 1021–1028. https://doi.org/10.1038/s41562-020-0918-6

Lindquist, K. A. (2017). The role of language in emotion: Existing evidence and future directions. *Current Opinion in Psychology*, *17*, 135–139. https://doi.org/10.1016/j.copsyc.2017.07.006

Lindquist, K. A. (2021). Language and Emotion: Introduction to the Special Issue. *Affective Science*, *2*(2), 91–98. https://doi.org/10.1007/s42761-021-00049-7

Lindquist, K. A., & Barrett, L. F. (2008). Constructing Emotion: The Experience of Fear as a Conceptual Act. *Psychological Science*, *19*(9), 898–903. https://doi.org/10.1111/j.1467-9280.2008.02174.x

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00444

LoBue, V., & Adolph, K. E. (2019). Fear in Infancy: Lessons From Snakes, Spiders, Heights, and Strangers. *Developmental Psychology*, *55*(9), 1889–1907. https://doi.org/10.1037/dev0000675

LoBue, V., Pérez-Edgar, K., & Buss, K. A. (2019). Introduction: Emotional Development, Past, and Present. In V. LoBue, K. Pérez-Edgar, & K. A. Buss (Eds.), *Handbook of Emotional Development* (pp. 1–3). Springer International Publishing. https://doi.org/10.1007/978-3-030-17332-6_1

Loria, S. (2018). *Textblob Documentation*. 73.

Lund, T. C., Sidhu, D. M., & Pexman, P. M. (2019). Sensitivity to emotion information in children's lexical processing. *Cognition*, *190*, 61–71. https://doi.org/10.1016/j.cognition.2019.04.017

Lutz, C. (1982). The domain of emotion words on Ifaluk. *American Ethnologist*, *9*(1), 113–128. https://doi.org/10.1525/ae.1982.9.1.02a00070

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00226

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs, Vol. 1, 3rd ed* (pp. xi, 366). Lawrence Erlbaum Associates Publishers.

Majid, A. (2012). Current Emotion Research in the Language Sciences. *Emotion Review*, *4*(4), 432–443. https://doi.org/10.1177/1754073912445827

Majid, A. (2022). The cultural landscape of emotions. *Science*, *377*(6602), 161–161. https://doi.org/10.1126/science.abq8098

Majid, A. (2023). Establishing psychological universals. *Nature Reviews Psychology*, *2*(4), 199–200. https://doi.org/10.1038/s44159-023-00169-w

Mak, M. H. C., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2023). Episodic memory and sleep are involved in the maintenance of context-specific lexical information. *Journal of Experimental Psychology. General*, *152*(11), 3087–3115. https://doi.org/10.1037/xge0001435

Mak, M. H. C., Hsiao, Y., & Nation, K. (2021). Anchoring and contextual variation in the early stages of incidental word learning during reading. *Journal of Memory and Language*, *118*, 104203. https://doi.org/10.1016/j.jml.2020.104203

Mak, M. H. C., O'Hagan, A., Horner, A. J., & Gaskell, M. G. (2023). A registered report testing the effect of sleep on Deese-Roediger-McDermott false memory: Greater lure and veridical recall but fewer intrusions after sleep. *Royal Society Open Science*, *10*(12), 220595. https://doi.org/10.1098/rsos.220595

Makowski, D. (2018). The psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science. *Journal of Open Source Software*, *3*(22), 470. https://doi.org/10.21105/joss.00470

Mar, R. A., & Oatley, K. (2008). The Function of Fiction is the Abstraction and Simulation of Social Experience. *Perspectives on Psychological Science*, *3*(3), 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, *40*(5), 694–712. https://doi.org/10.1016/j.jrp.2005.08.002

Mar, R. A., & Rain, M. (2015). Narrative Fiction and Expository Nonfiction Differentially Predict Verbal Ability. *Scientific Studies of Reading*, *19*(6), 419–433. https://doi.org/10.1080/10888438.2015.1069296

Marks, E. H., Walker, R. S. W., Ojalehto, H., Bedard-Gilligan, M. A., & Zoellner, L. A. (2019). Affect Labeling to Facilitate Inhibitory Learning: Clinical Considerations. *Cognitive and Behavioral Practice*, *26*(1), 201–213. https://doi.org/10.1016/j.cbpra.2018.05.001

Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*(3), 273–284. https://doi.org/10.1111/j.1467-9817.2008.00371.x

Matlin, M. W., & Stang, D. J. (1978). *The Pollyanna principle. Selectivity in language, memory, and thought.* Cambridge, MA: Schenkman.

McCarron, S. P., Murphy, V., & Nation, K. (2024). *An "Author Fluency Task": Semantic fluency as predictor of L2 vocabulary knowledge*. OSF. https://doi.org/10.31219/osf.io/dsx7w

McCutchen, D. (2006). Cognitive Factors in the Development of Children's Writing. In *Handbook of writing research* (pp. 115–130). The Guilford Press.

Millard, S. P. (2013). Hypothesis Tests. In S. P. Millard (Ed.), *EnvStats: An R Package for Environmental Statistics* (pp. 149–173). Springer. https://doi.org/10.1007/978-1-4614-8456-1_7

Mohammad, S. (2018a). Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 174–184. https://doi.org/10.18653/v1/P18-1017

Mohammad, S. (2018b, May). Word Affect Intensities. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. LREC 2018, Miyazaki, Japan. https://aclanthology.org/L18-1027

Mohammad, S. M., & Turney, P. D. (2010). Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 26–34. https://aclanthology.org/W10-0204

Mohammad, S. M., & Turney, P. D. (2013). *Crowdsourcing a Word-Emotion Association Lexicon* (arXiv:1308.6297). arXiv. https://doi.org/10.48550/arXiv.1308.6297

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The Words Children Hear: Picture Books and the Statistics for Language Learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning Word Meanings From Context During Normal Reading. *American Educational Research Journal*, *24*(2), 237–270. https://doi.org/10.3102/00028312024002237

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, *20*(2), 233–253. https://doi.org/10.2307/747758

Nation, K., Dawson, N. J., & Hsiao, Y. (2022). Book Language and Its Implications for Children's Language, Literacy, and Development. *Current Directions in Psychological Science*, *31*(4), 375–380. https://doi.org/10.1177/09637214221103264

Nencheva, M. L., Tamir, D. I., & Lew-Williams, C. (2023). Caregiver speech predicts the emergence of children's emotion vocabulary. *Child Development*. https://doi.org/10.1111/cdev.13897

Ng, B. C., Cui, C., & Cavallaro, F. (2019). The annotated lexicon of chinese emotion words. WORD, *65*(2), 73–92. https://doi.org/10.1080/00437956.2019.1599543

Noble, C., Cameron-Faulkner, T., Jessop, A., Coates, A., Sawyer, H., Taylor-Ims, R., & Rowland, C. F. (2020). The Impact of Interactive Shared Book Reading on Children's Language Skills: A Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research: JSLHR*, *63*(6), 1878–1897. https://doi.org/10.1044/2020_JSLHR-19-00288

Nook, E. C. (2021). Emotion Differentiation and Youth Mental Health: Current Understanding and Open Questions. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.700298

Nook, E. C. (2023). The Promise of Affective Language for Identifying and Intervening on Psychopathology. *Affective Science*, *4*(3), 517–521. https://doi.org/10.1007/s42761-023-00199-w

Nook, E. C., Hull, T. D., Nock, M. K., & Somerville, L. H. (2022). Linguistic measures of psychological distance track symptom levels and treatment outcomes in a large set of psychotherapy transcripts. *Proceedings of the National Academy of Sciences*, *119*(13), e2114737119. https://doi.org/10.1073/pnas.2114737119

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773–792. https://doi.org/10.1037/emo0000609

Ogren, M., & Sandhofer, C. M. (2021). Emotion words in early childhood: A language transcript analysis. *Cognitive Development, 60*, 101122. https://doi.org/10.1016/j.cogdev.2021.101122

O'Kearney, R., & Dadds, M. (2004). Developmental and gender differences in the language for emotions across the adolescent years. *Cognition and Emotion, 18*, 913–938. https://doi.org/10.1080/02699930341000356

Ong, D. C., Wu, Z., Zhi-Xuan, T., Reddan, M., Kahhale, I., Mattek, A., & Zaki, J. (2021). Modeling emotion in complex stories: The Stanford Emotional Narratives Dataset. *IEEE Transactions on Affective Computing*, *12*(3), 579–594. https://doi.org/10.1109/taffc.2019.2955949

Ornaghi, V., & Grazzani, I. (2013). The relationship between emotional-state language and emotion understanding: A study with school-age children. *Cognition and Emotion, 27*(2), 356–366. https://doi.org/10.1080/02699931.2012.711745

Ortony, A., Turner, T. J., & Antos, S. J. (1983). A puzzle about affect and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 725–729. https://doi.org/10.1037/0278-7393.9.4.725

Pagán, A., & Nation, K. (2019). Learning Words Via Reading: Contextual Diversity, Spacing, and Retrieval Effects in Adults. *Cognitive Science*, *43*(1). https://doi.org/10.1111/cogs.12705

Pauligk, S., Kotz, S. A., & Kanske, P. (2019). Differential Impact of Emotion on Semantic Processing of Abstract and Concrete Words: ERP and fMRI Evidence. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-50755-3

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language. use: Our words, our selves. *Annual Review of Psychology, 54*, 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Piantadosi, S. T. (2014). Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic Bulletin & Review*, *21*(5), 1112–1130. https://doi.org/10.3758/s13423-014-0585-6

Pinto, G., Tarchi, C., & Bigozzi, L. (2015). The relationship between oral and written narratives: A three-year longitudinal study of narrative cohesion, coherence, and structure. *British Journal of Educational Psychology*, *85*(4), 551–569. https://doi.org/10.1111/bjep.12091

Plaisted, H., Waite, P., & Creswell, C. (2022). Optimising exposure for adolescents with public speaking anxiety: Affect labelling or positive coping statements? *Behaviour Research and Therapy*, *148*, 103997. https://doi.org/10.1016/j.brat.2021.103997

Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental Science*, *21*(2), e12549. https://doi.org/10.1111/desc.12549

Ponari, M., Norbury, C. F., & Vigliocco, G. (2020). The role of emotional valence in learning novel abstract concepts. *Developmental Psychology*, *56*(10), 1855–1865. https://doi.org/10.1037/dev0001091

Pratto, F., & John, O. P. (1991a). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037//0022-3514.61.3.380

Pratto, F., & John, O. P. (1991b). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Ratner, C. (1989). A Social Constructionist Critique of The Naturalistic Theory of Emotion. *The Journal of Mind and Behavior*, *10*(3), 211–230.

Recchia, G., & Louwerse, M. M. (2015). Reproducing affective norms with lexical co-occurrence statistics: Predicting valence, arousal, and dominance. *Quarterly Journal of Experimental Psychology*, *68*(8), 1584–1598. https://doi.org/10.1080/17470218.2014.941296

Reggin, L. D., Muraki, E. J., & Pexman, P. M. (2021). Development of Abstract Word Knowledge. *Frontiers in Psychology*, *12*, 686478. https://doi.org/10.3389/fpsyg.2021.686478

Richards, B. (1987). Type/Token Ratios: What do they really tell us? *Journal of Child Language*, *14*(2), 201–209. https://doi.org/10.1017/S0305000900012885

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, *11*(3), 235–257. https://doi.org/10.1080/10888430701344306

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology*, *21*, 901–908. https://doi.org/10.1037/0012-1649.21.5.901

Rimehaug, T., & Kårstad, S. B. (2022). Communication and Emotional Vocabulary; Relevance for Mental Health Among School-Age Youths. *Frontiers in Psychology*, *13*, 847412. https://doi.org/10.3389/fpsyg.2022.847412

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Rosa, E., Salom, R., & Perea, M. (2022). Contextual diversity favors the learning of new words in children regardless of their comprehension skills. *Journal of Experimental Child Psychology*, *214*, 105312. https://doi.org/10.1016/j.jecp.2021.105312

Ruba, A. L., Pollak, S. D., & Saffran, J. R. (2022). Acquiring Complex Communicative Systems: Statistical Learning of Language and Emotion. *Topics in Cognitive Science*, *14*(3), 432–450. https://doi.org/10.1111/tops.12612

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295x.110.1.145

Russell, J. A., & Widen, S. C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*(1), 30–52. https://doi.org/10.1111/1467-9507.00185

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., & Kirkham, N. Z. (2018). Infant Statistical Learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Schwering, S. C., Ghaffari-Nikou, N. M., Zhao, F., Niedenthal, P. M., & MacDonald, M. C. (2021). Exploring the Relationship Between Fiction Reading and Emotion Recognition. *Affective Science*, *2*(2), 178–186. https://doi.org/10.1007/s42761-021-00034-0

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–104. https://doi.org/10.1016/j.biopsycho.2008.03.010

Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 783–792. https://doi.org/10.1037/a0027209

Seidenberg, M. S., & MacDonald, M. C. (2018). The Impact of Language Experience on Language and Reading: A Statistical Learning Approach. *Topics in Language Disorders*, *38*(1), 66. https://doi.org/10.1097/TLD.0000000000000144

Shablack, H., & Lindquist, K. A. (2019). The role of language in emotional development. In *Handbook of emotional development* (pp. 451–478). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-17332-6_18

Shaoul, C., & Westbury, C. (2013). *A reduced redundancy USENET corpus (2005-2011) Edmonton, AB: University of Alberta. Downloaded from*. https://www.psych.ualberta.ca/~westburylab/downloads/usenetcorpus.download.html

Shearer, N. J., Gillespie, A. N., Olds, T. S., Mensah, F. K., Edwards, B., Fernando, J. W., Wang, Y., Wake, M., & Lycett, K. (2021). Insights into adolescent well-being from computerised analysis of written language. *Acta Paediatrica (Oslo, Norway: 1992)*, *110*(6), 1880–1889. https://doi.org/10.1111/apa.15813

Silk, J. S., Siegle, G. J., Whalen, D. J., Ostapenko, L. J., Ladouceur, C. D., & Dahl, R. E. (2009). Pubertal Changes in Emotional Information Processing: Pupillary, Behavioral, and Subjective Evidence during Emotional Word Identification. *Development and Psychopathology*, *21*(1), 7–26. https://doi.org/10.1017/S0954579409000029

Siller, M., Swanson, M. R., Serlin, G., & George, A. (2014). Internal State Language in the Storybook Narratives of Children with and without Autism Spectrum Disorder: Investigating Relations to Theory of Mind Abilities. *Research in Autism Spectrum Disorders*, *8*(5), 589–596. https://doi.org/10.1016/j.rasd.2014.02.002

Silveira, B., Silva, H. S., Murai, F., & da Silva, A. P. C. (2021). Predicting user emotional tone in mental disorder online communities. *Future Generation Computer Systems*, *125*, 641–651. https://doi.org/10.1016/j.future.2021.07.014

Singh, J., Singh, G., & Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. *Human-Centric Computing and Information Sciences*, *7*(1), 32. https://doi.org/10.1186/s13673-017-0116-3

Snefjella, B., & Kuperman, V. (2016). It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition*, *156*, 135–146. https://doi.org/10.1016/j.cognition.2016.07.010

Snefjella, B., Lana, N., & Kuperman, V. (2020). How emotion is learned: Semantic learning of novel words in emotional contexts. *Journal of Memory and Language*, *115*, 104171. https://doi.org/10.1016/j.jml.2020.104171

Stainthorp, R. (1997). A Children's Author Recognition Test: A Useful Tool in Reading Research. *Journal of Research in Reading*, *20*(2), 148–158. https://doi.org/10.1111/1467-9817.00027

Stanovich, K. E., & West, R. F. (1989). Exposure to Print and Orthographic Processing. *Reading Research Quarterly*, *24*(4), 402–433. https://doi.org/10.2307/747605

Storm, C., & Storm, T. (1987). A Taxonomic Study of the Vocabulary of Emotions. *Journal of Personality and Social Psychology*, *53*(4), 805–816.

Streubel, B., Gunzenhauser, C., Grosse, G., & Saalbach, H. (2020). Emotion-specific vocabulary and its contribution to emotion understanding in 4- to 9-year-old children. *Journal of Experimental Child Psychology*, *193*, 104790. https://doi.org/10.1016/j.jecp.2019.104790

Talukdar, W., & Biswas, A. (2023). Improving Large Language Model (LLM) fidelity through context-aware grounding: A systematic approach to reliability and veracity. *World Journal of Advanced Engineering Technology and Sciences*, *10*(2), 283–296. https://doi.org/10.30574/wjaets.2023.10.2.0317

Thompson, R. A. (1991). Emotional regulation and emotional development. *Educational Psychology Review*, *3*, 269–307. https://doi.org/10.1007/BF01319934

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *TOWRE-2 Test of Word Reading Efficiency*. New York: Pearson.

Torre, J. B., & Lieberman, M. D. (2018). Putting Feelings Into Words: Affect Labeling as Implicit Emotion Regulation. *Emotion Review*, *10*(2), 116–124. https://doi.org/10.1177/1754073917742706

Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2020). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*, *195*, 103810. https://doi.org/10.1016/j.cognition.2018.09.013

Unkelbach, C., Alves, H., & Koch, A. (2020). Chapter Three - Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 62, pp. 115–187). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.005

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology*, *95*(1), 36–49. https://doi.org/10.1037/0022-3514.95.1.36

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valdivia-Moreno, C. A., Sasse, S. F., Lambert, H., McLaughlin, K. A., Somerville, L., & Nook, E. (2023). *Emotion word production develops in tandem with general verbal fluency and reveals key dimensions organizing emotion concepts*. OSF. https://doi.org/10.31219/osf.io/urf2w

van Bergen, E., Hart, S. A., Latvala, A., Vuoksimaa, E., Tolvanen, A., & Torppa, M. (2023). Literacy skills seem to fuel literacy enjoyment, rather than vice versa. *Developmental Science*, *26*(3), e13325. https://doi.org/10.1111/desc.13325

Van Goozen, S., & Frijda, N. H. (1993). Emotion words used in six European countries. *European Journal of Social Psychology*, *23*(1), 89–95. https://doi.org/10.1002/ejsp.2420230108

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading*, *29*(4), 367–382. https://doi.org/10.1111/j.1467-9817.2006.00301.x

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, *24*(7), 1767–1777. https://doi.org/10.1093/cercor/bht025

Vigliocco, G., Ponari, M., & Norbury, C. (2018). Learning and Processing Abstract Words and Concepts: Insights From Typical and Atypical Development. *Topics in Cognitive Science*, *10*(3), 533–549. https://doi.org/10.1111/tops.12347

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion*, *28*(4), 737–746. https://doi.org/10.1080/02699931.2013.851068

Voeten, C. C. (2023). *Using 'buildmer' to automatically find & compare maximal (mixed) models*. https://cran.r-project.org/web/packages/buildmer/vignettes/buildmer.html

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Warriner, A. B., & Kuperman, V. (2015). Affective biases in English are bi-dimensional. *Cognition and Emotion*, *29*(7), 1147–1167. https://doi.org/10.1080/02699931.2014.968098

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Weinstein, S. M., Mermelstein, R. J., Hankin, B. L., Hedeker, D., & Flay, B. R. (2007). Longitudinal Patterns of Daily Affect and Global Mood During Adolescence. *Journal of Research on*

Adolescence: The Official Journal of the Society for Research on Adolescence, *17*(3), 587–600. https://doi.org/10.1111/j.1532-7795.2007.00536.x

Werlen, E., Imhof, C., Benites, F., & Bergamin, P. B. (2019, September). *The reader's feeling and text-based emotions: The relationship between subjective self-reports, lexical ratings, and sentiment analysis*. SwissText 2019, Winterthur, 18 - 19 June 2019. https://doi.org/10.21256/zhaw-18859

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the Strange Stories: Revealing Mentalizing Impairments in Autism. *Child Development*, *80*(4), 1097–1117. https://doi.org/10.1111/j.1467-8624.2009.01319.x

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*, 114–128. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*, 291–312. https://doi.org/10.1016/j.cogdev.2008.01.002

Wierzbicka, A. (2009). Language and Metalanguage: Key Issues in Emotion Research. *Emotion Review*, *1*(1), 3–14. https://doi.org/10.1177/1754073908097175

Wild, K., Kilgarriff, A., & Tugwell, D. (2013). The Oxford Children's Corpus: Using a Children's Corpus in Lexicography1. *International Journal of Lexicography*, *26*(2), 190–218. https://doi.org/10.1093/ijl/ecs017

Wilmot, A., Hasking, P., Leitão, S., Hill, E., & Boyes, M. (2023). Understanding Mental Health in Developmental Dyslexia: A Scoping Review. *International Journal of Environmental Research and Public Health*, *20*(2), Article 2. https://doi.org/10.3390/ijerph20021653

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (Second edition). CRC Press/Taylor & Francis Group.

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*, *11*(9), 400–406. https://doi.org/10.1016/j.tics.2007.08.002

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Yang, J. S., Rosvold, C., & Bernstein Ratner, N. (2022). Measurement of Lexical Diversity in Children's Spoken Language: Computational and Conceptual Considerations. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.905789

Yao, Z., Yu, D., Wang, L., Zhu, X., Guo, J., & Wang, Z. (2016). Effects of valence and arousal on emotional word processing are modulated by concreteness: Behavioral and ERP evidence from a lexical decision task. *International Journal of Psychophysiology*, *110*, 231–242. https://doi.org/10.1016/j.ijpsycho.2016.07.499

Zagata, E., Kearns, D., Truckenmiller, A. J., & Zhao, Z. (2023). Using the Features of Written Compositions to Understand Reading Comprehension. *Reading Research Quarterly*. https://doi.org/10.1002/rrq.503

Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: A narrative review. *Npj Digital Medicine*, *5*(1), 1–13. https://doi.org/10.1038/s41746-022-00589-7

Zhang, Y., & Gu, Y. (2023). *A recipient design in multimodal language on TV: A comparison of child-directed and adult-directed broadcasting* [Proceedings paper]. 45th Annual Conference of the Cognitive Science Society. In: Proceedings of the Annual Meeting of the Cognitive Science Society. (Pp. Pp. 2869-2879). Cognitive Science Society: Seattle, WA, USA. (2023); Cognitive Science Society. https://escholarship.org/uc/item/17k7h7m6

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Appendices

Note on Appendix:

The appendices include information that was directly referenced but not included in the main text, such as full materials for experimental stimuli, additional methodological details, and supplementary analyses.

Additional materials associated with the thesis, such as preregistration documents, data files, and analysis scripts, are not attached here; they are available on each study's OSF page, as noted at the beginning of the thesis.

Appendix 1

I made a preliminary attempt to create a comprehensive emotion word list at the start of my DPhil. While this list was ultimately not used for corpus analysis in Chapter 2, it was referenced in Chapter 6 to assess children's emotion fluency. I briefly describe below the method for creating the emotion word list.

The preliminary lists of potential emotion words were compiled from a variety of sources, including websites, published papers, and the Oxford Children's Reading Corpus (see Table A-1). For online sources, I searched for the term "emotion words" using Google and selected the six most frequently visited websites that provided lists containing at least 100 emotion words. Additionally, emotion word lists were drawn from three existing studies on emotion vocabulary (Baron-Cohen et al., 2010; Ridgeway et al., 1985; Storm & Storm, 1987). Recognising that prior research had not extensively examined emotion words in the context of reading and written language, I also incorporated data from the Oxford Children's Reading Corpus. Using the lemma "feel" as a keyword, a frequency list of the 1000 words immediately following 'feel' was generated from the SketchEngine (https://www.sketchengine.eu/). All words on the list had a minimum frequency of 2. The list was then manually refined by excluding non-emotion-related terms, including stop words, modifying adverbs (e.g., "unusually," "extremely"), purely physiological sensations (e.g., "sore," "tight"), adjectives in comparative form (e.g., "better," "easier"), and other irrelevant terms, ultimately yielding a set of 419 emotion words. The distinct terms from these different sources were combined to produce a final list of 1198 emotion words.

Table A-1. Different sources for creating the emotion word list.

List	Source	Category	Number of words

1	https://www.enchantedlearning.com/wordlist/emotions.shtml	Online website	266
2	https://www.verywellfamily.com/feelings-words-from-a-to-z-2086647	Online website	353
3	https://www.centervention.com/list-of-emotions-135-words-that-express-feelings/	Online website	135
4	https://www.hoffmaninstitute.org/wp-content/uploads/Practices-FeelingsSensations.pdf	Online website	211
5	https://www.berkeleywellbeing.com/list-of-emotions.html	Online website	271
6	https://karlamclaren.com/emotional-vocabulary-page/	Online website	280
7	Ridgeway et al. (1985)	Paper	126
8	Storm and Storm (1987)	Paper	72
9	Baron-Cohen et al. (2010)	Paper	336
10	Oxford Children's Corpus - Reading 2019	Children's books	419

Appendix 2

This section is associated with Chapter 2. It first presents the graphs illustrating the age distributions of tokens in child-directed speech and in the picture book corpus (Figure A-1). It then reproduces the 336 emotion words in Baron-Cohen et al. (2010), along with the item level frequency statistics in different children's language corpora (Table A-2).

Figure A-1. Age distributions of tokens in child-directed speech and in the picture book corpus.

(a) child-directed speech (b) picture book corpus

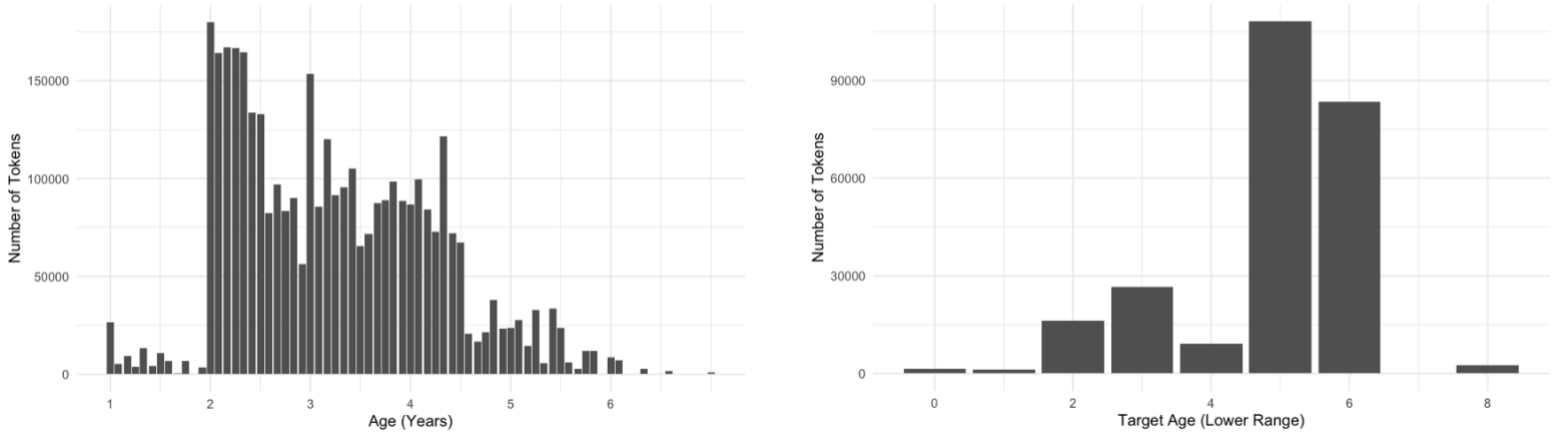


Table A-2. Item level statistics for the 336 emotion-description words in different children's language corpora.

Note:

Column A lists the 336 emotion-description words used in Baron-Cohen et al. (2010).

Column B-G reports the raw frequency of the emotion word in each of the 6 corpora.

Columns H-M reports the Zipf value for each emotion word in each corpus, following van Heuven et al. (2014).

The formula is: Zipf = log10((word_frequency_count + 1) / (corpus_token_count in millions + corpus_type_count in millions)) + 3.0

Word	Freq_ Spoken	Freq_ Picture book	Freq_ CBeebies	Freq_ CBBC	Freq_ Reading	Freq_ Writing	Zipf_ Spoken	Zipf_ Picture book	Zipf_ CBeebies	Zipf_ CBBC	Zipf_ Reading	Zipf_ Writing
sure	2268	139	4824	9916	28119	11317	5.77	5.63	5.91	5.86	5.64	5.37
sorry	2147	78	4781	10614	11626	10321	5.74	5.38	5.91	5.89	5.26	5.33
love	1766	126	8028	11640	22005	11761	5.66	5.58	6.14	5.93	5.54	5.39
hurt	1717	22	541	1525	7603	5471	5.65	4.84	4.96	5.05	5.08	5.06
mean	1421	91	2221	8271	18345	10329	5.56	5.44	5.58	5.78	5.46	5.33
cold	1330	66	1228	1989	13728	14239	5.54	5.31	5.32	5.16	5.33	5.47
broken	1156	18	511	1053	8500	5298	5.47	4.76	4.94	4.89	5.12	5.04
happy	990	142	3510	5291	13055	20192	5.41	5.64	5.78	5.59	5.31	5.62
lost	990	80	1436	3243	15464	13477	5.41	5.39	5.39	5.38	5.38	5.45
cross	675	51	417	1034	5628	2335	5.24	5.20	4.85	4.88	4.95	4.69
wanted	653	102	1349	3940	19671	27582	5.23	5.49	5.36	5.46	5.49	5.76
sad	524	89	790	932	5579	8185	5.13	5.43	5.13	4.83	4.94	5.23

warm	523	60	1238	1177	7715	6358	5.13	5.27	5.32	4.94	5.08	5.12
wonder	505	38	2949	1517	8270	4021	5.12	5.07	5.70	5.05	5.11	4.92
close	450	56	1044	6276	17211	6680	5.07	5.24	5.25	5.66	5.43	5.14
kind	419	81	1152	6312	19375	9741	5.03	5.39	5.29	5.66	5.48	5.31
giving	406	29	335	1722	6740	2475	5.02	4.96	4.76	5.10	5.02	4.71
difficult	395	15	311	1739	5915	902	5.01	4.68	4.73	5.10	4.97	4.27
fine	362	47	1359	4278	12789	6565	4.97	5.16	5.36	5.50	5.30	5.14
scared	328	61	777	2742	3821	17809	4.93	5.27	5.12	5.30	4.78	5.57
lying	328	44	131	680	6466	5085	4.93	5.13	4.35	4.70	5.01	5.02
listening	260	25	295	859	3955	1911	4.83	4.90	4.70	4.80	4.79	4.60
thinking	217	59	896	2987	11610	11065	4.75	5.26	5.18	5.34	5.26	5.36
afraid	192	57	538	1730	9816	2542	4.70	5.24	4.96	5.10	5.19	4.72
gentle	178	8	204	316	3442	1283	4.66	4.43	4.54	4.37	4.73	4.43
liked	170	69	480	1126	6574	5378	4.64	5.33	4.91	4.92	5.01	5.05
upset	167	16	317	803	2003	2821	4.64	4.71	4.73	4.77	4.50	4.77
excited	164	38	737	2292	3370	8312	4.63	5.07	5.10	5.22	4.72	5.24
calm	152	9	329	1339	3975	3948	4.60	4.48	4.75	4.99	4.79	4.91
nasty	149	23	172	723	1388	1304	4.59	4.86	4.47	4.72	4.34	4.43
moved	144	35	298	763	10515	6441	4.57	5.04	4.71	4.75	5.22	5.13
safe	141	33	1419	2646	8454	8315	4.56	5.01	5.38	5.29	5.12	5.24
interested	136	8	100	743	3802	754	4.55	4.43	4.24	4.74	4.77	4.20
pleased	131	51	404	989	5385	1544	4.53	5.20	4.84	4.86	4.93	4.51
frightened	129	23	162	357	4423	3110	4.53	4.86	4.44	4.42	4.84	4.81
bothered	105	2	24	283	758	602	4.44	3.96	3.63	4.32	4.07	4.10
glad	103	24	836	1199	11263	1642	4.43	4.88	5.15	4.94	5.25	4.53
comfortable	99	17	127	520	3184	834	4.41	4.74	4.34	4.58	4.70	4.24
surprised	98	26	156	689	5455	3640	4.41	4.91	4.43	4.70	4.93	4.88
shy	91	8	169	471	1140	1160	4.38	4.43	4.46	4.54	4.25	4.38

bored	88	14	189	430	1069	2910	4.36	4.66	4.51	4.50	4.22	4.78
angry	87	39	121	1342	6201	5755	4.36	5.08	4.32	4.99	4.99	5.08
worried	85	22	373	1378	3719	7124	4.35	4.84	4.80	5.00	4.77	5.17
friendly	74	39	261	567	3084	2852	4.29	5.08	4.65	4.62	4.68	4.77
stupid	74	14	34	2167	4028	2598	4.29	4.66	3.78	5.20	4.80	4.73
joking	71	5	141	512	588	495	4.27	4.26	4.38	4.57	3.96	4.01
trapped	70	2	121	1547	1687	3896	4.26	3.96	4.32	5.05	4.42	4.91
jolly	65	31	143	297	1930	587	4.23	4.99	4.39	4.34	4.48	4.09
needed	61	22	516	1299	7748	11097	4.20	4.84	4.94	4.98	5.08	5.36
hate	52	21	70	1527	3608	2655	4.14	4.82	4.08	5.05	4.75	4.74
low	48	29	451	969	10754	2143	4.10	4.96	4.89	4.85	5.23	4.65
helpful	46	0	239	197	850	799	4.08	3.48	4.61	4.16	4.12	4.22
lonely	45	23	187	308	2550	4187	4.07	4.86	4.51	4.35	4.60	4.94
brave	42	24	867	1108	4085	4121	4.04	4.88	5.17	4.91	4.81	4.93
wondering	42	10	240	567	3682	4699	4.04	4.52	4.61	4.62	4.76	4.99
miserable	40	7	26	167	2243	1445	4.02	4.38	3.66	4.09	4.55	4.48
merry	40	7	174	220	3064	293	4.02	4.38	4.47	4.21	4.68	3.79
hopeless	39	1	49	136	858	475	4.01	3.78	3.93	4.00	4.13	4.00
tender	39	1	9	51	2030	135	4.01	3.78	3.23	3.58	4.50	3.45
grumpy	36	18	212	178	275	1387	3.98	4.76	4.56	4.12	3.64	4.46
touched	35	12	49	251	4503	3152	3.97	4.59	3.93	4.27	4.85	4.82
serious	34	17	161	2055	4584	1126	3.96	4.74	4.44	5.18	4.86	4.37
fierce	31	23	48	328	2259	1299	3.92	4.86	3.92	4.38	4.55	4.43
annoyed	29	5	19	226	1119	1666	3.89	4.26	3.53	4.22	4.24	4.54
certain	28	7	76	673	10236	750	3.87	4.38	4.12	4.69	5.20	4.19
nosy	28	0	12	33	82	123	3.87	3.48	3.34	3.40	3.11	3.41
keen	27	7	110	626	1963	342	3.86	4.38	4.28	4.66	4.49	3.85
confused	27	0	66	470	2349	7193	3.86	3.48	4.06	4.54	4.57	5.18

desperate	25	3	23	482	2445	1484	3.83	4.08	3.61	4.55	4.58	4.49
teasing	24	5	19	42	460	207	3.81	4.26	3.53	3.50	3.86	3.64
settled	23	19	143	258	4657	1309	3.79	4.78	4.39	4.28	4.86	4.44
concentrating	23	1	32	159	338	181	3.79	3.78	3.75	4.07	3.72	3.58
cruel	22	5	17	344	2333	1186	3.77	4.26	3.49	4.40	4.56	4.39
impressed	22	3	70	926	1231	577	3.77	4.08	4.08	4.83	4.29	4.08
determined	21	9	22	468	3740	1987	3.75	4.48	3.59	4.54	4.77	4.62
pleasure	20	12	205	524	5206	365	3.73	4.59	4.54	4.58	4.91	3.88
uncomfortable	20	5	10	124	1454	632	3.73	4.26	3.27	3.96	4.36	4.12
terror	20	1	8	151	2852	1670	3.73	3.78	3.19	4.05	4.65	4.54
complaining	20	0	20	99	527	283	3.73	3.48	3.55	3.86	3.92	3.77
useless	19	3	43	365	1759	672	3.71	4.08	3.87	4.43	4.44	4.15
bold	19	3	33	181	1572	596	3.71	4.08	3.76	4.12	4.39	4.09
nervous	18	11	135	2343	2556	3280	3.69	4.56	4.36	5.23	4.60	4.83
panic	18	7	196	573	1550	2498	3.69	4.38	4.53	4.62	4.39	4.72
generous	18	0	17	118	1342	331	3.69	3.48	3.49	3.94	4.32	3.84
fond	17	9	15	91	3022	276	3.67	4.48	3.44	3.83	4.68	3.76
favor	17	0	0	7	2263	17	3.67	3.48	2.23	2.77	4.55	2.57
obsessed	17	0	2	137	168	293	3.67	3.48	2.71	4.00	3.42	3.79
proud	16	24	239	1670	4342	2772	3.64	4.88	4.61	5.09	4.83	4.76
jealous	15	6	35	361	1228	965	3.62	4.33	3.79	4.42	4.28	4.30
fascinated	15	1	4	30	580	238	3.62	3.78	2.93	3.36	3.96	3.70
knowing	14	22	45	443	5349	3558	3.59	4.84	3.89	4.51	4.92	4.87
delighted	13	16	70	201	2478	1054	3.56	4.71	4.08	4.17	4.59	4.34
attacked	13	3	9	230	2101	1815	3.56	4.08	3.23	4.23	4.52	4.58
bitter	13	1	153	82	1998	666	3.56	3.78	4.42	3.78	4.50	4.14
playful	13	0	23	84	294	244	3.56	3.48	3.61	3.79	3.66	3.71
fear	12	39	241	1338	10284	6937	3.53	5.08	4.61	4.99	5.21	5.16

relief	12	1	42	599	4525	1996	3.53	3.78	3.86	4.64	4.85	4.62
amazed	11	6	57	221	1313	2572	3.49	4.33	3.99	4.21	4.31	4.73
pity	11	3	67	134	2802	152	3.49	4.08	4.06	3.99	4.64	3.50
cheerful	11	1	34	38	2029	465	3.49	3.78	3.78	3.46	4.50	3.99
hurried	10	32	37	5	3798	1369	3.45	5.00	3.81	2.64	4.77	4.45
unhappy	10	10	46	120	2014	458	3.45	4.52	3.90	3.95	4.50	3.98
choosing	10	6	121	205	676	148	3.45	4.33	4.32	4.18	4.03	3.49
foolish	10	2	11	105	1919	307	3.45	3.96	3.31	3.89	4.48	3.81
grief	10	1	23	113	1910	233	3.45	3.78	3.61	3.92	4.48	3.69
alert	10	0	133	316	919	520	3.45	3.48	4.36	4.37	4.16	4.04
jumpy	9	11	118	50	107	90	3.41	4.56	4.31	3.57	3.23	3.28
ashamed	9	7	11	107	2119	385	3.41	4.38	3.31	3.90	4.52	3.90
embarrassed	9	4	24	245	998	734	3.41	4.18	3.63	4.26	4.19	4.18
impatient	9	1	9	23	1109	227	3.41	3.78	3.23	3.24	4.24	3.68
crafty	9	1	28	122	198	66	3.41	3.78	3.69	3.95	3.49	3.14
cheated	9	0	18	90	304	121	3.41	3.48	3.51	3.82	3.68	3.40
crushed	9	0	32	157	1084	677	3.41	3.48	3.75	4.06	4.23	4.15
grateful	8	7	29	139	1961	702	3.37	4.38	3.71	4.01	4.49	4.17
peaceful	8	5	116	148	1625	1610	3.37	4.26	4.30	4.04	4.41	4.53
shocked	8	5	23	281	1758	5310	3.37	4.26	3.61	4.31	4.44	5.04
lively	8	3	46	90	1333	188	3.37	4.08	3.90	3.82	4.32	3.59
sneaky	8	1	88	314	110	522	3.37	3.78	4.18	4.36	3.24	4.04
dread	8	1	83	99	1188	394	3.37	3.78	4.16	3.86	4.27	3.91
moody	7	46	0	62	297	286	3.31	5.15	2.23	3.66	3.67	3.78
disappointed	7	11	102	534	1928	989	3.31	4.56	4.24	4.59	4.48	4.31
satisfied	7	5	9	65	2575	239	3.31	4.26	3.23	3.68	4.61	3.70
battered	7	3	5	54	1055	728	3.31	4.08	3.01	3.60	4.22	4.18
mischievous	7	0	6	75	388	661	3.31	3.48	3.08	3.75	3.78	4.14

secure	7	0	21	286	1488	265	3.31	3.48	3.57	4.32	4.37	3.74
cheered	6	24	54	33	942	1940	3.26	4.88	3.97	3.40	4.17	4.61
joy	6	19	131	388	4855	2975	3.26	4.78	4.35	4.45	4.88	4.79
convinced	6	3	4	229	1746	427	3.26	4.08	2.93	4.23	4.44	3.95
involved	6	1	13	894	2535	422	3.26	3.78	3.38	4.82	4.60	3.94
outgoing	6	0	0	30	51	30	3.26	3.48	2.23	3.36	2.91	2.81
frustrated	6	0	1	129	331	605	3.26	3.48	2.53	3.98	3.72	4.10
contrary	6	0	4	48	1350	18	3.26	3.48	2.93	3.55	4.33	2.60
terrified	5	14	15	273	1397	4058	3.19	4.66	3.44	4.30	4.34	4.93
cunning	5	5	17	245	1408	522	3.19	4.26	3.49	4.26	4.34	4.04
deciding	5	2	35	94	398	277	3.19	3.96	3.79	3.84	3.80	3.76
eager	5	2	11	96	2498	823	3.19	3.96	3.31	3.85	4.59	4.23
injured	5	1	20	339	1416	1320	3.19	3.78	3.55	4.40	4.35	4.44
guilty	5	1	6	310	1480	729	3.19	3.78	3.08	4.36	4.37	4.18
definite	5	1	4	115	491	57	3.19	3.78	2.93	3.93	3.89	3.08
vain	5	1	2	54	2196	191	3.19	3.78	2.71	3.60	4.54	3.60
whining	5	0	124	25	273	220	3.19	3.48	4.33	3.28	3.63	3.66
curious	4	12	31	360	3635	2227	3.11	4.59	3.74	4.42	4.76	4.67
ignored	4	7	6	63	1394	1679	3.11	4.38	3.08	3.67	4.34	4.54
sensitive	4	3	39	338	970	129	3.11	4.08	3.83	4.39	4.18	3.43
depressed	4	1	2	42	557	505	3.11	3.78	2.71	3.50	3.94	4.02
innocent	4	0	3	264	2143	1461	3.11	3.48	2.83	4.29	4.53	4.48
detached	4	0	0	13	309	93	3.11	3.48	2.23	3.01	3.69	3.29
threatening	4	0	2	98	932	494	3.11	3.48	2.71	3.86	4.16	4.01
restless	4	0	4	28	946	194	3.11	3.48	2.93	3.33	4.17	3.61
agony	4	0	7	62	1261	607	3.11	3.48	3.13	3.66	4.30	4.10
controlled	4	0	17	289	1083	454	3.11	3.48	3.49	4.33	4.23	3.98
gloomy	3	8	13	21	1051	3776	3.01	4.43	3.38	3.21	4.22	4.90

puzzled	3	6	19	9	2001	1308	3.01	4.33	3.53	2.86	4.50	4.44
grave	3	3	1	132	3336	1272	3.01	4.08	2.53	3.99	4.72	4.42
tempted	3	3	0	59	710	216	3.01	4.08	2.23	3.64	4.05	3.65
responsible	3	1	8	431	1414	284	3.01	3.78	3.19	4.50	4.35	3.77
aware	3	1	3	243	3147	635	3.01	3.78	2.83	4.25	4.69	4.12
tearful	3	0	1	20	166	116	3.01	3.48	2.53	3.19	3.42	3.39
encouraging	3	0	1	97	572	113	3.01	3.48	2.53	3.86	3.95	3.38
demanding	3	0	2	144	431	157	3.01	3.48	2.71	4.03	3.83	3.52
content	3	0	2	66	3921	232	3.01	3.48	2.71	3.69	4.79	3.69
cooperative	3	0	0	7	55	2	3.01	3.48	2.23	2.77	2.94	1.80
baffled	3	0	3	29	339	166	3.01	3.48	2.83	3.34	3.73	3.54
hollow	2	16	62	65	2665	495	2.89	4.71	4.03	3.68	4.62	4.01
daring	2	12	25	182	1101	608	2.89	4.59	3.65	4.13	4.24	4.10
furious	2	9	22	184	1454	1630	2.89	4.48	3.59	4.13	4.36	4.53
begging	2	9	3	110	727	638	2.89	4.48	2.83	3.91	4.06	4.12
dignified	2	6	4	13	517	17	2.89	4.33	2.93	3.01	3.91	2.57
amused	2	3	6	36	1224	113	2.89	4.08	3.08	3.43	4.28	3.38
modest	2	3	0	39	730	43	2.89	4.08	2.23	3.47	4.06	2.96
thankful	2	2	3	25	902	321	2.89	3.96	2.83	3.28	4.15	3.83
hopeful	2	2	15	48	589	221	2.89	3.96	3.44	3.55	3.97	3.66
disturbed	2	1	12	63	1426	352	2.89	3.78	3.34	3.67	4.35	3.87
spiteful	2	1	0	7	186	115	2.89	3.78	2.23	2.77	3.47	3.38
thrilled	2	1	18	82	473	388	2.89	3.78	3.51	3.78	3.87	3.91
disgrace	2	1	8	68	667	141	2.89	3.78	3.19	3.70	4.02	3.47
distress	2	1	18	80	1280	241	2.89	3.78	3.51	3.77	4.30	3.70
flattered	2	1	0	24	304	16	2.89	3.78	2.23	3.26	3.68	2.55
wishful	2	0	0	0	42	10	2.89	3.48	2.23	1.86	2.83	2.36
heated	2	0	28	69	678	149	2.89	3.48	3.69	3.71	4.03	3.49

comforting	2	0	0	19	570	387	2.89	3.48	2.23	3.17	3.95	3.91
protective	2	0	24	229	515	282	2.89	3.48	3.63	4.23	3.91	3.77
displeased	2	0	0	7	262	13	2.89	3.48	2.23	2.77	3.61	2.46
aggressive	2	0	0	260	431	271	2.89	3.48	2.23	4.28	3.83	3.75
numb	2	0	4	36	309	292	2.89	3.48	2.93	3.43	3.69	3.79
bashful	2	0	0	10	179	8	2.89	3.48	2.23	2.91	3.45	2.27
understanding	1	9	12	133	2080	214	2.71	4.48	3.34	3.99	4.51	3.65
weary	1	5	8	19	1951	354	2.71	4.26	3.19	3.17	4.49	3.87
distracted	1	4	18	257	736	980	2.71	4.18	3.51	4.28	4.06	4.31
helpless	1	4	3	52	1254	608	2.71	4.18	2.83	3.59	4.29	4.10
longing	1	4	6	6	1138	271	2.71	4.18	3.08	2.71	4.25	3.75
assured	1	3	4	50	1503	134	2.71	4.08	2.93	3.57	4.37	3.45
relaxed	1	2	27	295	869	617	2.71	3.96	3.68	4.34	4.13	4.11
homesick	1	2	15	30	252	96	2.71	3.96	3.44	3.36	3.60	3.30
respect	1	2	14	521	2938	355	2.71	3.96	3.41	4.58	4.66	3.87
regret	1	2	16	201	1355	699	2.71	3.96	3.46	4.17	4.33	4.16
sulky	1	1	13	7	270	25	2.71	3.78	3.38	2.77	3.63	2.73
grouchy	1	1	2	11	44	43	2.71	3.78	2.71	2.94	2.85	2.96
threatened	1	1	4	139	1384	371	2.71	3.78	2.93	4.01	4.34	3.89
desire	1	1	4	83	3141	229	2.71	3.78	2.93	3.79	4.69	3.68
daydreaming	1	0	5	16	58	339	2.71	3.48	3.01	3.09	2.97	3.85
enjoyment	1	0	2	20	733	71	2.71	3.48	2.71	3.19	4.06	3.18
questioning	1	0	1	42	537	303	2.71	3.48	2.53	3.50	3.93	3.80
argumentative	1	0	0	14	36	11	2.71	3.48	2.23	3.04	2.76	2.40
accepting	1	0	0	13	413	60	2.71	3.48	2.23	3.01	3.81	3.10
irritated	1	0	1	19	423	198	2.71	3.48	2.53	3.17	3.82	3.62
encouraged	1	0	1	69	1086	224	2.71	3.48	2.53	3.71	4.23	3.67
misunderstood	1	0	1	35	156	37	2.71	3.48	2.53	3.42	3.39	2.90

heartache	1	0	0	9	76	9	2.71	3.48	2.23	2.86	3.08	2.32
defeated	1	0	6	181	1326	1984	2.71	3.48	3.08	4.12	4.32	4.62
cowardly	1	0	0	13	632	80	2.71	3.48	2.23	3.01	4.00	3.23
considerate	1	0	3	16	187	50	2.71	3.48	2.83	3.09	3.47	3.03
affectionate	1	0	1	17	723	48	2.71	3.48	2.53	3.12	4.05	3.01
mistreated	1	0	0	6	26	30	2.71	3.48	2.23	2.71	2.63	2.81
judging	1	0	31	198	466	118	2.71	3.48	3.74	4.16	3.86	3.39
disinterested	1	0	0	0	109	8	2.71	3.48	2.23	1.86	3.24	2.27
accusing	1	0	0	30	201	31	2.71	3.48	2.23	3.36	3.50	2.82
enthusiastic	1	0	11	99	506	170	2.71	3.48	3.31	3.86	3.90	3.55
critical	1	0	1	134	1054	39	2.71	3.48	2.53	3.99	4.22	2.92
rejected	1	0	0	52	514	121	2.71	3.48	2.23	3.59	3.91	3.40
sickened	1	0	0	2	113	10	2.71	3.48	2.23	2.34	3.25	2.36
uptight	1	0	3	20	22	4	2.71	3.48	2.83	3.19	2.56	2.02
misjudged	1	0	1	16	51	12	2.71	3.48	2.53	3.09	2.91	2.43
overpowering	1	0	0	14	157	57	2.71	3.48	2.23	3.04	3.39	3.08
overwhelmed	1	0	0	63	470	474	2.71	3.48	2.23	3.67	3.87	3.99
bliss	1	0	8	18	288	116	2.71	3.48	3.19	3.14	3.66	3.39
weepy	0	14	3	5	18	10	2.41	4.66	2.83	2.64	2.47	2.36
startled	0	10	13	16	1994	1113	2.41	4.52	3.38	3.09	4.49	4.37
anxious	0	10	5	106	3742	1168	2.41	4.52	3.01	3.89	4.77	4.39
distant	0	5	31	81	3216	921	2.41	4.26	3.74	3.78	4.70	4.28
astonished	0	4	7	9	1585	925	2.41	4.18	3.13	2.86	4.39	4.28
loyal	0	4	12	123	748	388	2.41	4.18	3.34	3.96	4.07	3.91
caring	0	3	24	148	640	1053	2.41	4.08	3.63	4.04	4.00	4.34
thoughtful	0	3	26	58	1024	223	2.41	4.08	3.66	3.64	4.21	3.67
cranky	0	3	6	29	78	95	2.41	4.08	3.08	3.34	3.09	3.30
trusting	0	2	2	40	425	65	2.41	3.96	2.71	3.48	3.82	3.14

disgust	0	2	2	17	906	319	2.41	3.96	2.71	3.12	4.15	3.82
troubled	0	2	1	46	1652	147	2.41	3.96	2.53	3.54	4.41	3.49
alarmed	0	2	4	46	1054	367	2.41	3.96	2.93	3.54	4.22	3.88
overjoyed	0	2	3	10	270	494	2.41	3.96	2.83	2.91	3.63	4.01
disbelief	0	2	2	6	331	685	2.41	3.96	2.71	2.71	3.72	4.15
dazed	0	2	3	8	553	394	2.41	3.96	2.83	2.82	3.94	3.91
appalled	0	2	0	15	339	55	2.41	3.96	2.23	3.07	3.73	3.07
unfriendly	0	1	3	6	186	64	2.41	3.78	2.83	2.71	3.47	3.13
dreamy	0	1	51	23	296	155	2.41	3.78	3.95	3.24	3.67	3.51
unsure	0	1	6	51	394	456	2.41	3.78	3.08	3.58	3.79	3.98
horrified	0	1	1	17	660	991	2.41	3.78	2.53	3.12	4.01	4.31
forgiving	0	1	1	9	140	41	2.41	3.78	2.53	2.86	3.34	2.94
glum	0	1	35	25	128	174	2.41	3.78	3.79	3.28	3.31	3.56
shaken	0	1	9	43	972	221	2.41	3.78	3.23	3.51	4.18	3.66
watchful	0	1	3	32	383	32	2.41	3.78	2.83	3.38	3.78	2.84
disagreeable	0	1	0	0	668	4	2.41	3.78	2.23	1.86	4.02	2.02
worthless	0	1	4	52	322	162	2.41	3.78	2.93	3.59	3.70	3.53
insulted	0	1	6	36	294	50	2.41	3.78	3.08	3.43	3.66	3.03
tense	0	1	3	348	699	270	2.41	3.78	2.83	4.41	4.04	3.75
offended	0	1	1	46	734	96	2.41	3.78	2.53	3.54	4.06	3.30
harsh	0	1	5	219	1183	528	2.41	3.78	3.01	4.21	4.27	4.04
boastful	0	1	0	1	80	50	2.41	3.78	2.23	2.17	3.10	3.03
courageous	0	1	2	51	258	406	2.41	3.78	2.71	3.58	3.61	3.93
needy	0	1	1	21	248	30	2.41	3.78	2.53	3.21	3.59	2.81
overcome	0	1	5	235	1156	257	2.41	3.78	3.01	4.24	4.26	3.73
cautious	0	1	3	107	515	286	2.41	3.78	2.83	3.90	3.91	3.78
forbidding	0	1	0	10	326	35	2.41	3.78	2.23	2.91	3.71	2.87
absorbed	0	1	0	35	1223	148	2.41	3.78	2.23	3.42	4.28	3.49

discouraged	0	1	0	20	363	12	2.41	3.78	2.23	3.19	3.76	2.43
inspired	0	1	10	302	1048	227	2.41	3.78	3.27	4.35	4.22	3.68
dismayed	0	1	0	1	230	29	2.41	3.78	2.23	2.17	3.56	2.80
devoted	0	1	1	43	1327	40	2.41	3.78	2.53	3.51	4.32	2.93
fed-up	0	0	0	0	16	0	2.41	3.48	2.23	1.86	2.43	1.32
dislike	0	0	1	30	629	81	2.41	3.48	2.53	3.36	3.99	3.23
bad-tempered	0	0	0	0	89	0	2.41	3.48	2.23	1.86	3.15	1.32
uncaring	0	0	0	3	23	16	2.41	3.48	2.23	2.47	2.57	2.55
agreeable	0	0	0	7	1159	3	2.41	3.48	2.23	2.77	4.26	1.92
brokenhearted	0	0	0	2	3	3	2.41	3.48	2.23	2.34	1.80	1.92
appreciated	0	0	3	51	389	119	2.41	3.48	2.83	3.58	3.79	3.40
heartbroken	0	0	1	17	92	366	2.41	3.48	2.53	3.12	3.16	3.88
adoring	0	0	3	35	98	28	2.41	3.48	2.83	3.42	3.19	2.78
calculating	0	0	3	8	216	11	2.41	3.48	2.83	2.82	3.53	2.40
lovesick	0	0	0	7	13	2	2.41	3.48	2.23	2.77	2.34	1.80
disbelieving	0	0	0	2	58	14	2.41	3.48	2.23	2.34	2.97	2.49
challenging	0	0	2	279	294	190	2.41	3.48	2.71	4.31	3.66	3.60
apologetic	0	0	1	1	173	33	2.41	3.48	2.53	2.17	3.44	2.85
distrust	0	0	0	0	297	4	2.41	3.48	2.23	1.86	3.67	2.02
forceful	0	0	0	24	81	43	2.41	3.48	2.23	3.26	3.11	2.96
easy-going	0	0	0	0	77	0	2.41	3.48	2.23	1.86	3.09	1.32
approving	0	0	0	1	186	4	2.41	3.48	2.23	2.17	3.47	2.02
despair	0	0	16	59	2049	691	2.41	3.48	3.46	3.64	4.51	4.16
uncertain	0	0	0	27	815	93	2.41	3.48	2.23	3.31	4.11	3.29
murderous	0	0	0	9	262	212	2.41	3.48	2.23	2.86	3.61	3.65
hassled	0	0	0	6	10	8	2.41	3.48	2.23	2.71	2.24	2.27
insulting	0	0	1	34	170	39	2.41	3.48	2.53	3.41	3.43	2.92
attracted	0	0	5	83	1128	224	2.41	3.48	3.01	3.79	4.25	3.67

clueless	0	0	0	38	48	199	2.41	3.48	2.23	3.46	2.88	3.62
dissatisfied	0	0	0	3	166	8	2.41	3.48	2.23	2.47	3.42	2.27
doubtful	0	0	6	10	727	29	2.41	3.48	3.08	2.91	4.06	2.80
focused	0	0	6	372	946	472	2.41	3.48	3.08	4.44	4.17	3.99
disapproval	0	0	0	0	200	17	2.41	3.48	2.23	1.86	3.50	2.57
bewildered	0	0	0	7	867	435	2.41	3.48	2.23	2.77	4.13	3.96
detesting	0	0	0	0	2	1	2.41	3.48	2.23	1.86	1.67	1.62
outraged	0	0	3	15	256	117	2.41	3.48	2.83	3.07	3.60	3.39
humiliated	0	0	0	38	153	96	2.41	3.48	2.23	3.46	3.38	3.30
neglected	0	0	0	20	749	131	2.41	3.48	2.23	3.19	4.07	3.44
mocking	0	0	0	13	402	114	2.41	3.48	2.23	3.01	3.80	3.38
discomfort	0	0	0	17	299	32	2.41	3.48	2.23	3.12	3.67	2.84
disrespectful	0	0	0	25	108	67	2.41	3.48	2.23	3.28	3.23	3.15
charmed	0	0	5	17	434	15	2.41	3.48	3.01	3.12	3.83	2.52
fired-up	0	0	0	0	1	0	2.41	3.48	2.23	1.86	1.50	1.32
defensive	0	0	0	86	439	39	2.41	3.48	2.23	3.80	3.84	2.92
distaste	0	0	0	1	137	9	2.41	3.48	2.23	2.17	3.33	2.32
carefree	0	0	0	14	64	39	2.41	3.48	2.23	3.04	3.01	2.92
betrayed	0	0	1	50	981	210	2.41	3.48	2.53	3.57	4.19	3.64
charitable	0	0	0	17	223	9	2.41	3.48	2.23	3.12	3.54	2.32
discontented	0	0	1	0	217	5	2.41	3.48	2.53	1.86	3.53	2.10
astounded	0	0	6	9	202	154	2.41	3.48	3.08	2.86	3.50	3.51
anticipating	0	0	2	7	147	29	2.41	3.48	2.71	2.77	3.36	2.80
humored	0	0	0	0	10	1	2.41	3.48	2.23	1.86	2.24	1.62
overpowered	0	0	0	7	242	75	2.41	3.48	2.23	2.77	3.58	3.20
craving	0	0	1	14	153	51	2.41	3.48	2.53	3.04	3.38	3.03
compassionate	0	0	0	7	129	35	2.41	3.48	2.23	2.77	3.31	2.87
cocky	0	0	1	48	168	42	2.41	3.48	2.53	3.55	3.42	2.95

discouraging	0	0	0	0	85	4	2.41	3.48	2.23	1.86	3.13	2.02
devious	0	0	1	154	81	162	2.41	3.48	2.53	4.05	3.11	3.53
dominated	0	0	0	27	426	66	2.41	3.48	2.23	3.31	3.83	3.14
disheartened	0	0	0	17	86	69	2.41	3.48	2.23	3.12	3.13	3.16
dictating	0	0	0	0	28	2	2.41	3.48	2.23	1.86	2.66	1.80
conceited	0	0	0	8	201	12	2.41	3.48	2.23	2.82	3.50	2.43
assessing	0	0	0	18	93	6	2.41	3.48	2.23	3.14	3.17	2.16
dispirited	0	0	0	0	71	4	2.41	3.48	2.23	1.86	3.05	2.02
mystified	0	0	0	1	144	34	2.41	3.48	2.23	2.17	3.36	2.86
confronted	0	0	0	12	439	209	2.41	3.48	2.23	2.98	3.84	3.64
disregard	0	0	0	7	161	7	2.41	3.48	2.23	2.77	3.40	2.22
contempt	0	0	0	3	761	11	2.41	3.48	2.23	2.47	4.08	2.40
confrontational	0	0	0	3	7	0	2.41	3.48	2.23	2.47	2.10	1.32
commiserating	0	0	0	3	10	0	2.41	3.48	2.23	2.47	2.24	1.32
complacent	0	0	0	20	91	3	2.41	3.48	2.23	3.19	3.16	1.92
condescending	0	0	0	0	83	5	2.41	3.48	2.23	1.86	3.12	2.10

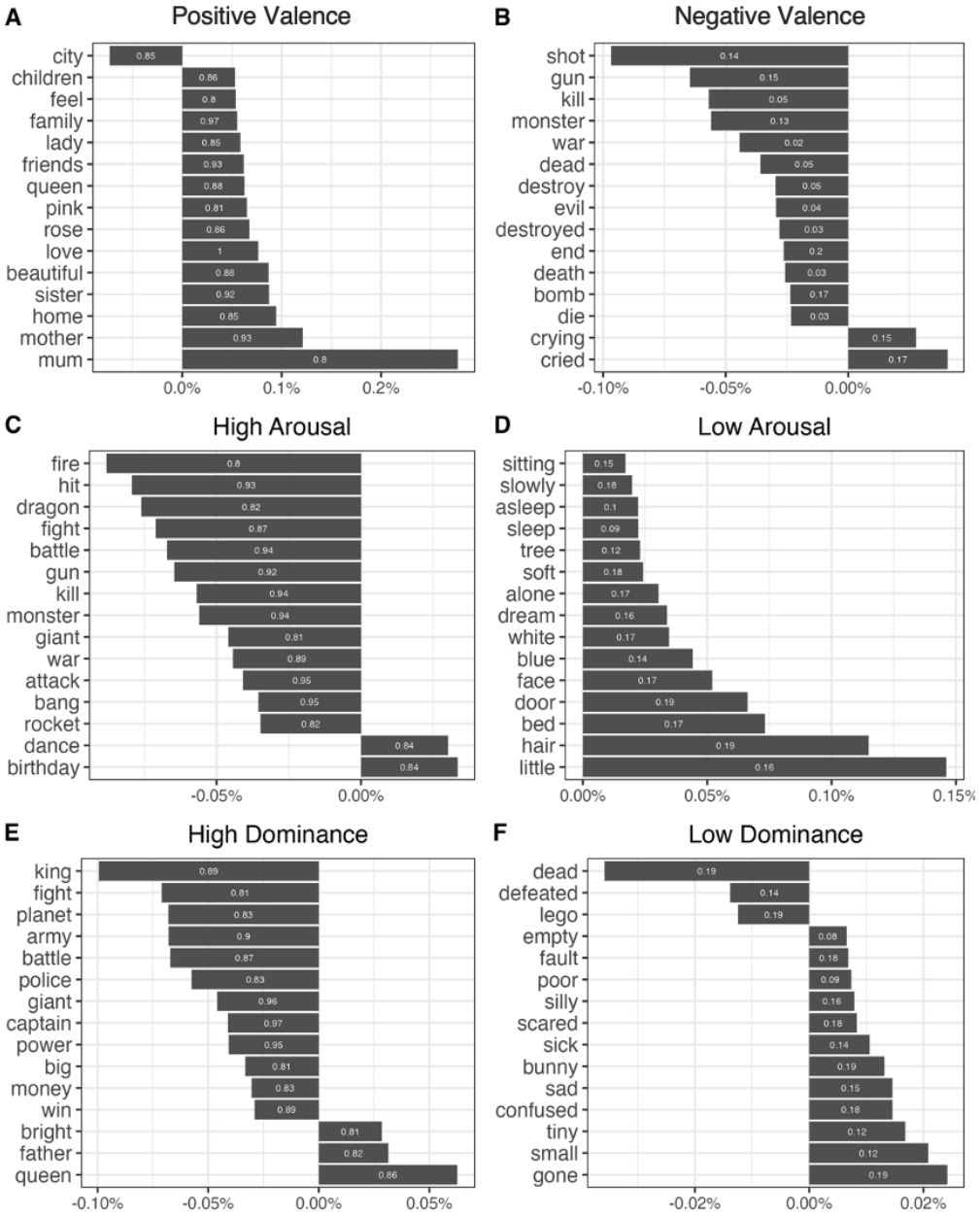
Appendix 3

This appendix is associated with Chapter 3. It gives some examples of boys' and girls' use of words and sentences in their writing.

3A. Examples of gender differences in the use of words

We extracted top words in various dimensions. These words either had very positive (larger than 0.8) or very negative (less than 0.2) scores for valence; very high (larger than 0.8) or very low (less than 0.2) scores for arousal and dominance. We also extracted words with very high (larger than 0.8) scores in the four emotion categories. We extracted top words which were used most differently by boys and girls, as shown in Figure A-2. X-axis shows the average percent difference in word usage, calculated as girl usage subtracted by boy usage. Positive percentage, or bars pointing to the right, therefore reflects greater usage of the word by girls. The graph only shows the difference in percentage usage, but not the actual percentage of usage. For example, among the top words with high valence scores of larger than 0.8, the word 'friends' has a valence score of 0.93 and was used 0.06% more in stories by girls (0.33%) than in stories by boys (0.27%). Among the words with low valence scores of less than 0.2, the word 'shot' has a valence score of 0.14 and was used 0.09% more in stories by boys (0.14%) than in stories by girls (0.05%).

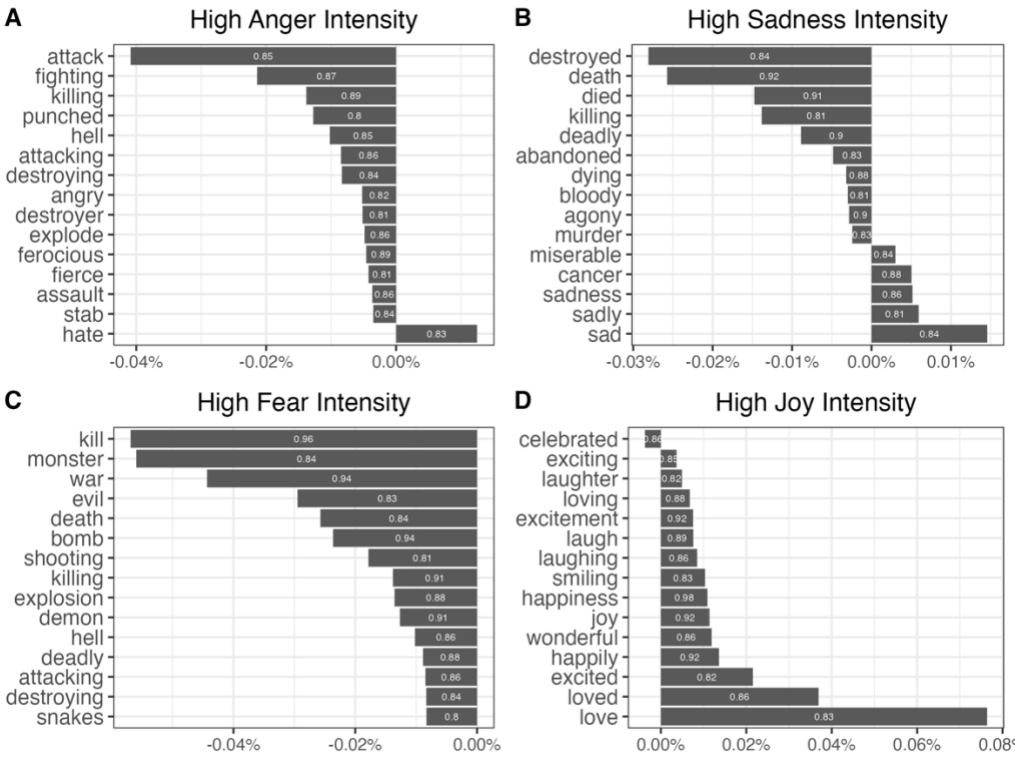
Figure A-2. Gender differences in word occurrences for words with (A) positive valence, (B) negative valence, (C) high arousal, (D) low arousal, (E) high dominance, and (F) low dominance scores.



Note. Y-axis shows 15 words that differ the most in usage between boys and girls in the respective dimension. Values inside the bars are the Valence, Arousal, and Dominance scores for the words on the respective dimensions.

Figure A-3 shows the 4 sets of 15 top words (scores > 0.8 for the 4 emotions categories) used most differently between boys and girls.

Figure A-3. Gender differences in word occurrences for words with (A) high anger intensity, (B) high sadness intensity, (C) high fear intensity, (D) high joy intensity.



Note. Y-axis shows 15 words that differ the most in usage between boys and girls in the respective emotion category. Values inside the bars are the intensity scores for the respective emotion category.

3B. Examples of gender differences in the use of sentences

"Once a upon a time there was a boy named Tom and a boy named Alex who were friends (actually brothers). They loved haunted things. They read books about ghosts, zombies, dragons and that sort of stuff. Here is where our story begins." (Boy, 7 years old)

“As they approached London they saw blazes of fire and sparks filling the air. Super Sheep tightened his trusty magic cape and was ready for anything that got in his way.” (Boy, 9 years old)

“I'm wondering if that note could be a sign of hope and refuge to get me out of this place, where the law is corrupted, people go out with intentions of death, blood fills the street. till we had a night curfew were people fight to try to be free.” (Boy, 13 years old)

“Luckily in Bobs backpack, he had a piece of rope, Evie grabbed it and, gave Bob a ginormous hug! Evie now feels happyish. She feels happy because, she was saved and now still alive.” (Girl, 7 years old)

“At the beach it was very sunny so they went in the ocean to swim, they also had a scrumtious picnic for lunch. They played on the beach a bit more and eventually went back to the caravan to have dinner and get settled for an early night in bed as they had another busy day ahead of them tomorow.” (Girl, 11 years old)

“One of things that makes my sadness go away is thinking of all the stories people make up about me, children look up at the shimmering sky and ask what I am, their parents reply that a man lives on me and he has a special name.” (Girl, 13 years old)

Appendix 4

4A. Target novel words and foils used in the adult experiment

Table A-3. Target novel words and foils used in the adult experiment.

Target novel words	Foils and distractors
rarive	geosive
venvic	cusive
depish	horive
kunable	shonbic
raxous	milatic
jemical	huxatic
quanful	thunish
binushy	tredish
supular	tumish
picial	gilable
higgive	gufable
roific	hixable
wapish	merous
sopable	chutous
domous	weerous
paxible	berical
priful	temical
shromy	vebical
jafular	shurful

hitial	glaful
thutive	croiful
talulic	scemowy
yeotish	chrabby
bulable	roithy
plarous	pimular
vorical	lonular
spomful	ferular
phrinty	dracial
vatular	phutial
cretial	jetial
	mibical
	genful
	mollive
	weacial
	hilish
	vuxable
	dolous
	cedular
	vepatic
	celoppy

4B. Narratives used in the adult experiment

Table A-4. Narratives used in the adult experiment.

Context Valence	Block	Blank Content
Positive	One	It was my birthday yesterday. I had a _____ cake together with some _____ friends.
		My friends held a party in their _____ garden last week. It was a _____ day with warm sunshine and a gentle breeze.
		I got myself a _____ cat! I love holding it in my arms. It is soft and _____.
		It was _____ to see my best friends again after so many years. We had a _____ chat along the river.
		I sat next to a _____ classmate today. She is really _____ and helped me a lot.
		I won a _____ prize in writing today. I love the feeling of writing _____ stories.
		This weekend, I had a _____ time at home. I watched a _____ movie and felt very relaxed.
		I helped my grandma clean her house today. It feels _____ to be helpful. We had a _____ time together.
		After working hard, I finally completed my homework. How _____. It feels _____ to be able to enjoy my life again.
		I am having a _____ time with my family in this beautiful weather. We enjoyed the _____ scenery.
Neutral		I was reading the cookbook: "First wash the ingredients using _____ water and pour some oil into a _____ pan."

		I went to the _____ library today. I first turned left at the _____ house.
		I booked a flight ticket this morning. I clicked on the _____ icon on the website, and I selected the _____ button.
		This _____ machine was newly produced by the company. It has a _____ cover and four wheels.
		I drove my _____ way home today. Everything was normal, nothing happened. I am used to driving on this _____ road.
		In the morning, I usually wake up on my _____ bed. I have cereal from my _____ bowl.
		I am at the _____ train station and reading the timetable. The _____ train will leave at 7 PM.
		I tried using this _____ kettle to boil water today. The _____ kettle was made of steel.
		I wanted to return a book. I went to the _____ building and turned right when I reached the _____ door.
		I went to see a _____ house today. It had several _____ rooms and a normal kitchen.
Negative		The _____ video games gave me a nightmare. The characters were fighting in a _____ way.
		It is so _____ that the _____ teacher only scolded me today, when many others were also fighting.
		I felt _____ when I saw the news about the _____ car accident. Several people were injured!
		I failed again after many _____ attempts at the maths challenge. I felt so _____.

		I had a _____ argument with my friend today. We could not agree. Their _____ words hurt my feelings.
		I was _____ today after working nonstop for hours. I felt _____ when going through endless homework.
		Someone said that I look _____, and they fear hanging out with a _____ monster like me.
		I fell down the stairs and injured myself. How _____! My legs are _____ now.
		This _____ place is giving me a headache. It is so _____ to hear these loud sounds.
		My plan today was ruined due to the _____ weather! The sky is dark and _____.
Positive	Two	On my birthday, my friends brought me so many _____ presents. I love them for being _____.
		I had a _____ picnic with my friends. It was a beautiful day, and we had a _____ time together.
		My cat made this _____ purr sound that made my heart melt. I love living with this _____ pet.
		I hope my _____ friends can come back to visit more often. We had a such a _____ time catching up today.
		It was _____ to meet a new classmate today. We had a _____ chat and hung out together.
		My teacher said she enjoyed the _____ story I wrote for the writing class. She said it was _____.
		I was very engaged with the _____ movie. I also enjoyed some _____ pizza while watching it.

		I feel _____ for helping my grandma. She thanked me for my _____ work.
		Finishing my work in time is _____. I will go to a _____ party tonight to celebrate.
		My family travelled together last week. We enjoyed the _____ mountains and the _____ sunrise.
Neutral		The cookbook says that "the next step is to place it into a _____ bowl. After you finish, wash the dish with _____ water."
		I saw a _____ post office on my way to the library. I walked straight and saw a _____ building on the right.
		The _____ button directed me to the next page. I then had to click the _____ box to confirm the booking.
		This machine has a _____ mark of the company name in the middle. There is also a _____ button on its cover.
		I drove the normal way home, seeing the _____ things as usual. It was just a _____ day anyway.
		In the morning, I go to school by hopping on a _____ bus. I always carry my _____ bag with me.
		I am at the train station. The ticket says that I should wait at the _____ platform and be ready to get on the _____ train.
		I boiled water with the _____ kettle today by putting some water into it. I switched it off after the _____ light flashed.
		I went to return a book. I opened the _____ door and put the book on the _____ bookcase.
		I do not know who lives in this _____ house. There is a _____ balcony facing the street.

Negative		I hate these kinds of _____ video games. The _____ scene still makes me shiver.
		The _____ principal only scolded me when others were also fighting. I cannot stand this at all. This is so _____.
		This accident has caused _____ damage to the cars. It has also resulted in _____ injuries to several people.
		I failed at my maths exam again. Am I really _____ at maths? Maybe I am just stupid and _____ at everything.
		I fought with my friend today. I feel so _____ and I hope to not have a _____ fight like this again.
		The _____ deadline for my homework is just a few days from now. I am so _____ and unprepared.
		Being judged by my _____ appearance makes me sad. I cannot choose how I look, even if it is _____.
		I fell down the stairs and my legs are _____. I should be more careful next time to avoid this _____ accident.
		I want to leave this _____ place at once. It is unbearable. Every second here makes me more _____.
		My mood is ruined by the _____ weather. I have lost motivation to work. I hope the _____ weather will be over soon.

4C. Norming and validation of stimuli

Prior to data collection, the valence of the stimuli was estimated using a *Bidirectional Encoder Representations from Transformers (BERT)* model, a transformer-based machine learning technique

for natural language processing (Devlin et al., 2019). *BERT* considers the context for each occurrence of a given word to generate word embedding representations for each word. For this validation of stimuli, we directly applied the model *"nlptown/bert-base-multilingual-uncased-sentiment"* from Hugging Face, which was pretrained on product reviews and fine-tuned for sentiment analysis. The output of the model was one of five sentiment classes, from negative to positive. To get a more nuanced sentiment score, we used the *softmax* function to obtain the probability of all five sentiment classes by exponentiating the model's output logits and normalising them so that they summed up to 1. From the resulting matrix, we assigned a weight to each sentiment class to obtain the resulting sentiment score for each story. The score was a scalar value that ranges from -1 (very negative) to 1 (very positive). The mean estimated sentiment was -0.41 (SD = 0.43) for the narratives in the negative context, -0.12 (SD = 0.40) in the neutral context, and 0.77 (SD = 0.13) in the positive context. Note that while BERT took into account the context to determine sentiment, it was not foolproof and could make wrong predictions. For example, the context about a fight with my friend was rated as 0.70 (very positive) by BERT, which deviated from its intended message. Nevertheless, one-way ANOVA showed that the three conditions were significantly different in their sentiment scores.

To ascertain that the stimuli achieved their intended valence effects, after data collection was completed, we conducted a post hoc analysis of the valence of the stimuli. The independent variable was the pre-determined valence of the narrative context. The dependent variable was the participants' judgment of the valence of those narratives. The same inclusion criteria applied for the validation study participants, who all reported to be native English speakers based in the UK area, have normal or corrected-to-normal vision, and no history of dyslexia or other language difficulties. They all provided consent before taking part.

Valence categorisation

Twenty participants (18 Females, 2 Males, M_{age} = 18.39, SD_{age} = 0.56, range = 18-19) read through the 60 narratives one at a time. They completed a valence categorisation task after reading each narrative to sort each narrative into either the "positive", "neutral", or "negative" valence category. Overall, the

mean agreement rate (participants sorting a narrative into the pre-determined category) was 96.75% (*SD* = 17.7%). No item had an agreement rate of less than 80%.

Valence rating

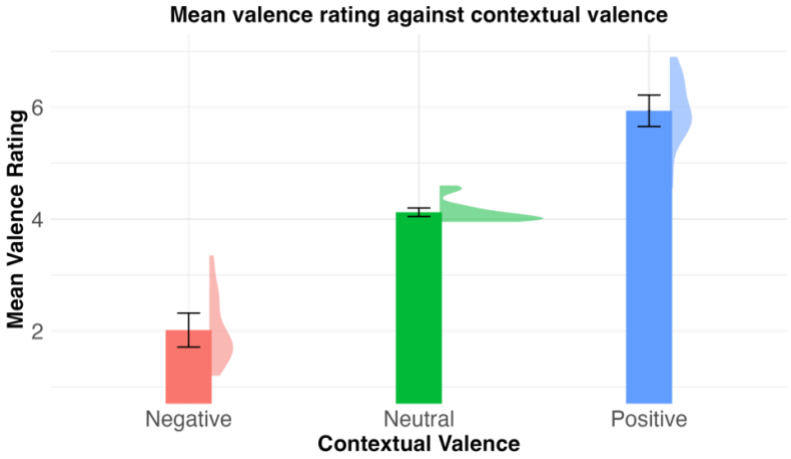
Another group of 19 participants (10 Females, 9 Males, M_{age} = 25.58, SD_{age} = 3.40, range = 19-30) did a valence rating task after reading each of the 60 narratives, where they were asked to rate on a Likert scale of 1-7, how positive/negative the narrative makes them feel. The descriptive statistics of their ratings (*M* and *SD*) are reported in the table and figure below. Participants rating of the valence of the narratives were consistent with their categorisation and *BERT* estimation.

We also calculated a correlation between *BERT* ratings and valence ratings. Pearson correlation shows that BERT ratings and valence ratings are strongly positively correlated, $r = 0.69$, $p < .001$.

Table A-5. Mean valence ratings of narrative stimuli by 19 participants who did not participate in the main experiment.

Valence	M valence rating	SD valence rating
Negative	2.02	0.60
Neutral	4.12	0.19
Positive	5.94	0.57

Figure A-4. Mean valence ratings of narrative stimuli by 19 participants who did not participate in the main experiment.



4D. Results of exit questionnaires

At the end of Session 1 participants were asked if they were aware of the study's aim. Analysis of the results showed that while 31 participants (41%) indicated yes, only 10 of them mentioned the word "context" in their responses and only one participant mentioned emotionality or valence of the context.

Appendix 5

Appendix 5 describes the stimuli used in the Study 2 of Chapter 5. These stimuli were selected and adapted from those used in Chapter 4.

5A. Target novel words and foils used in the child experiment

Table A-6. Target novel words and foils used in the child experiment.

Target novel words	Foils
cusive	gesive
roific	kanive
wapish	dobive
larous	chabic
breful	miatic
vebive	phutic
hurnic	jemish
tumish	vorish
domous	yotish
lenful	vatous
garive	futous
venvic	jafous
depish	thuful
merous	spoful
priful	criful
cusive	gesive

roific	kanive
wapish	dobive
larous	chabic
breful	miatic

5B. Narratives used in child experiment

Table A-7. Narratives used in the child experiment.

Context Valence	Block	Blank Content
Positive	One	It was my birthday yesterday. I had a _____ cake together with some _____ friends.
		I got myself a _____ cat. I love holding it in my arms. It is soft and _____.
		I sat next to a _____ classmate today. She is really _____ and helped me a lot.
		I won a _____ prize in writing today. I love the feeling of writing _____ stories.
		I am having a _____ time with my family in this beautiful weather. We enjoyed the _____ views.
Neutral		I was reading the cookbook. I need to wash the tomatoes using _____ water. Then I need to pour some oil into the _____ pan.

		I went to the _____ library today. To get there, I turned left at the _____ house.
		In the morning, I usually wake up on my _____ bed. I have cereal from my _____ bowl.
		I attended a _____ assembly this morning. During the _____ assembly, the headteacher played the piano.
		I wanted to return a book. I went to the _____ building. I turned right when I reached the _____ door.
Negative		The _____ video games gave me a nightmare. The way the characters were fighting was _____.
		My plan today was ruined because of the _____ weather. The sky is dark and _____.
		It was _____ that the _____ teacher only told me off today. Many others were also fighting.
		I failed again after many _____ attempts at the maths challenge. I felt _____ and wanted to cry.
		I had a _____ argument with my classmate today. Their _____ words hurt my feelings.
Positive	Two	It was my birthday yesterday. I had a _____ cake together with some _____ friends.
		I got myself a _____ cat. I love holding it in my arms. It is soft and _____.
		I sat next to a _____ classmate today. She is really _____ and helped me a lot.
		I won a _____ prize in writing today. I love the feeling of writing _____ stories.

		I am having a _____ time with my family in this beautiful weather. We enjoyed the _____ views.
Neutral		I was reading the cookbook. I need to wash the tomatoes using _____ water. Then I need to pour some oil into the _____ pan.
		I went to the _____ library today. To get there, I turned left at the _____ house.
		In the morning, I usually wake up on my _____ bed. I have cereal from my _____ bowl.
		I attended a _____ assembly this morning. During the _____ assembly, the headteacher played the piano.
Negative		I wanted to return a book. I went to the _____ building. I turned right when I reached the _____ door.
		The _____ video games gave me a nightmare. The way the characters were fighting was _____.
		My plan today was ruined because of the _____ weather. The sky is dark and _____.
		It was _____ that the _____ teacher only told me off today. Many others were also fighting.
		I failed again after many _____ attempts at the maths challenge. I felt _____ and wanted to cry.

Appendix 6

6A. Author Recognition Test items

Table A-8. Items in the Author Recognition Test.

Real authors	Foils
David Walliams	Catherine Harmer
J.K. Rowling	D Gaffan
Roald Dahl	David Popplewell
Dav Pilky	G Claridge
Dr Seuss	DA Allport
Jeff Kinney	Kia Nobre
Michael Morpurgo	M Treisman
Jamie Smart	S Millar
Julia Donaldson	M Rushworth
JRR Tolkien	Mansur Lalljee
Lewis Carroll	PE Bryant
Tony Ross	Paul Azzopardi
C.S. Lewis	Charles Spence
Liz Pichon	Edmund Rolls
Emma Carroll	Larry Weiskrantz
Philip Pullman	Nick Rawlins
Rick Riordan	Brian Rogers
Terry Deary	Jane Mellanby
Jacqueline Wilson	Kim Plunkett

Enid Blyton	GB Henning
Katherine Rundell	Catherine Harmer
Robin Stevens	D Gaffan
Shirley Hughes	David Popplewell
David McKee	G Claridge
John Boyne	DA Allport
Mick Inkpen	
Dick King-Smith	
Jill Murphy	
JM Barrie	
LM Montgomery	
David Walliams	
J.K. Rowling	
Roald Dahl	
Dav Pilky	
Dr Seuss	

6B. Further description of the animal list used to score verbal fluency

The primary objective of compiling reference lists for animal names and emotion words was to establish a broad catalogue that facilitates the initial phase of fluency scoring. This approach allows for quick identification, though each item was not individually verified. Any item from a child's production that did not appear in the list was flagged for manual review to assess its relevance.

The list of animal name, adapted in July 2024 from an online A-Z animal listing (https://a-z-animals.com/animals/), comprises 2870 names. The process for constructing the emotion word list is detailed further in Appendix 1, resulting in a total of 1198 emotion words.