
What words do children encounter when they read for pleasure?

The ability to read opens up worlds. Reading enables children to progress into post-primary education and provides the basis for lifelong learning and prosperity into adulthood. Importantly, the [journey](#) to becoming a skilled reader requires not only high-quality classroom instruction but also many years of practice through independent book reading.

We wanted to learn more about the vocabulary that children encounter when they read for pleasure. To do this, we analysed the words in 1200 books popular with British children aged 7–16. The original research article is [open access and free to download](#), and we summarise the key insights from this work below.

Books contain a vast number of words

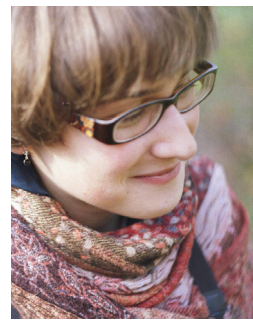
Reading seems so fast and automatic that sometimes people think that to be able to read we just memorise printed words. This idea has led to strategies that try to teach children to memorise the shapes of words. However, this type of rote learning is effortful and takes a long time. For instance, children in China need to memorise 2500 characters during primary school, and this takes around 9 hours per week for 6 years!

A working knowledge of 2000–3000 characters is enough to understand most modern texts in standard Chinese. Not so in English – the 1200 children’s books that we analysed contain over 100,000 different words! There is no way that a child could memorise so many printed words. That’s why *phonics is so powerful*. Without understanding the connections between letters and sounds, children won’t be able to break down the wide variety of words that they will encounter during independent reading.

Books contain many words that children may not know

We found that around 40% of words used in children’s books do not appear on BBC television programs aimed at children of the same age. Similarly, one fifth of words used in books for young people aged 13–16 are not encountered on BBC channels targeting adult audiences. The most common of these words include rare and sophisticated vocabulary – often of foreign origin – related to science (e.g. ‘meridian’, ‘homunculus’), arts (e.g. ‘aria’), history (e.g. ‘marquis’, ‘inquisitor’), politics (e.g. ‘communists’, ‘suffragists’, ‘abolitionist’, ‘legislature’) and religion (e.g. ‘quaker’, ‘missionary’). Typically, if a word is not in our spoken vocabulary, we can use context to infer meaning. However, if a book includes too many words that a child does not know, reading it will be a struggle.

The large numbers of unfamiliar words mean that *books present a unique opportunity for enhancing children’s vocabulary*. However, the other side of the coin is that, *for many children, reading is likely to pose a challenge from the earliest years of independent reading*.



**Maria
Korochkina**



**Kathy
Rastle**

What words do children encounter when they read for pleasure?

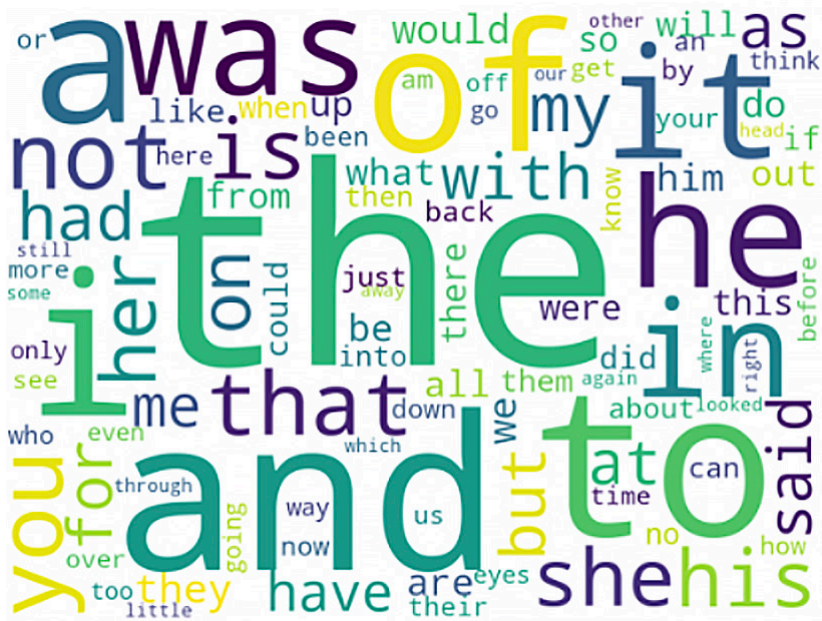


Figure 1: 100 most common words in the 1200 children's books we analysed.

and how it contributes to the overall meaning of the complex word. This body of knowledge is referred to as morphological knowledge.

The 1200 children's books that we analysed use thousands of morphologically complex words. Examples include words like 'inexpensively', 'unwinnable', 'unlawfulness', 'speechlessness' or 'outlandishly' – these words appear in books for children aged 10–12, but not in books for younger children. As skilled readers, we can easily understand these words even if we haven't encountered them before because each of these words is created by combining elements we already know (e.g. 'speech' + '-less' + '-ness'). However, these words will be very challenging to those who have not yet learned that 'in-' and 'un-' mean 'not', or that '-ness' denotes a noun and '-ly' an adverb. For these reasons, strong morphological knowledge is key to being able to read well.

Conclusions

It is widely accepted that reading ability is a strong predictor of how much children choose to read. Our work suggests that a failure to acquire good phonetic and morphological knowledge early in reading acquisition is likely to have a negative snowball effect on a child's reading habits. On the other hand, our analyses show that the books popular with British children today offer a wonderful opportunity to build vocabulary, particularly if children read widely.

Few words are used repeatedly in books

We found that the most common 100 words make up around 54% of the 1200 books that we analysed. Most of these words belong to a class of words called function words: these are words like 'do', 'and', 'not', 'but', or 'is', which are used to express relationships between other words. Every second word encountered in children's books is a function word, and children will quickly learn to recognise these words by sight. This form of sight-word recognition increases the speed of reading, but these words carry little meaning, and being able to read them quickly will not be enough to understand what a text is about.

To illustrate, consider a sentence from one of the books we analysed where all but the function words have been removed: "Then ... a ... her ... and she ... her ... and ... her." Can you guess what this is about? Now consider the original sentence: "Then a mischievous thought flashed across her eyes, and she pursed her lips together and pushed her tongue forward." This example shows why *being able to read the top 100 words effortlessly is not sufficient to read for meaning.*

New words are encountered in every book

It turns out that the vast majority of words in children's books are only encountered a few times and in a small subset of books. One consequence of this is that books written for children of the same age tend to vary greatly in

the words they use. This is particularly so for books aimed at younger primary school children. The words in these are less similar to one another compared to the words in books written for older children. Likewise, books for older children include many words that are not encountered in books for younger children. For instance, more than one third of words in books for children aged 10–12 are never used in books for younger children, and more than one third of words in books for young people aged 13–16 do not occur in books targeting older primary school children. This means that reading is likely to continue to pose a challenge as children grow older.

With book vocabulary being so intense, it is crucial to develop reading skills and motivation early on. And because different books use different words, it is important that children read widely.

Most new words have complex structure

We have said that the vocabulary in books is more sophisticated than the vocabulary on television and tends to get richer as children age. One way that we see this richness is through morphological complexity. Morphologically complex words are words that consist of several elements that are themselves meaningful: for example, the word 'mistrustfulness' consists of four elements, 'mis-', '-trust-', '-ful', and '-ness'. To understand the meanings of these words, a child needs to know what each individual component means

This article originally appeared on the [Rastle Lab blog](#).

Maria Korochkina (<https://mariakna.github.io/>) is a Postdoctoral Research Fellow in the Department of Psychology at Royal Holloway, University of London. Maria's research to date has focused on how we learn and consolidate new words, and on how word learning is influenced by the context in which new words are encountered.

Kathy Rastle [[@Kathy_Rastle](#) on X] is Professor of Psychology at Royal Holloway, University of London. She studies the neurocognitive mechanisms that underpin how people read and learn to read. She is very interested in using insights from research on reading to improve outcomes for children.