



Dr Maria Korochkina

## **The words children and young people encounter when they read for pleasure, and what it means for learning**

Research and Practice meeting  
27 June 2025



**Economic  
and Social  
Research Council**



# CYP-LEX

## The Children and Young People's Books Lexicon

*1,200 popular books, 400 books per age band*

7-9 years



10-12 years



13-16 years



# CYP-LEX

## The Children and Young People's Books Lexicon

*1,200 popular books, 400 books per age band*

7-9 years



10-12 years



13-16 years





# CYP-LEX

## The Children and Young People's Books Lexicon

*1,200 popular books, 400 books per age band*

7-9 years



10-12 years



13-16 years



# CYP-LEX

## The Children and Young People's Books Lexicon

*1,200 popular books, 400 books per age band*

7-9 years



10-12 years



13-16 years

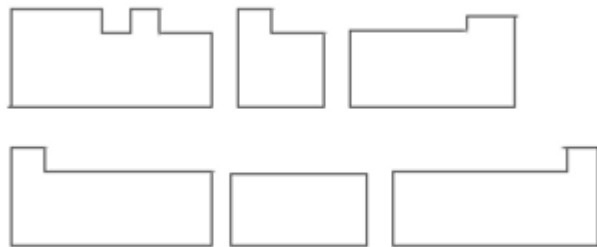


The **lightbulb** moments



# Books contain **vast** numbers of words

- Over **70 million words** and **over 100,000** *distinct* words in 1,200 books
- **50,000+** distinct words in the 7-9 age band alone



**This is not  
how we read.**



- Memorising words by rote is not an option
- Phonics is key!



# Many of these words will be **unfamiliar**

- **40%** of distinct words in books don't appear in children's TV programmes



- **21%** of distinct words in books for children aged 13-16 do not appear on adult TV programmes

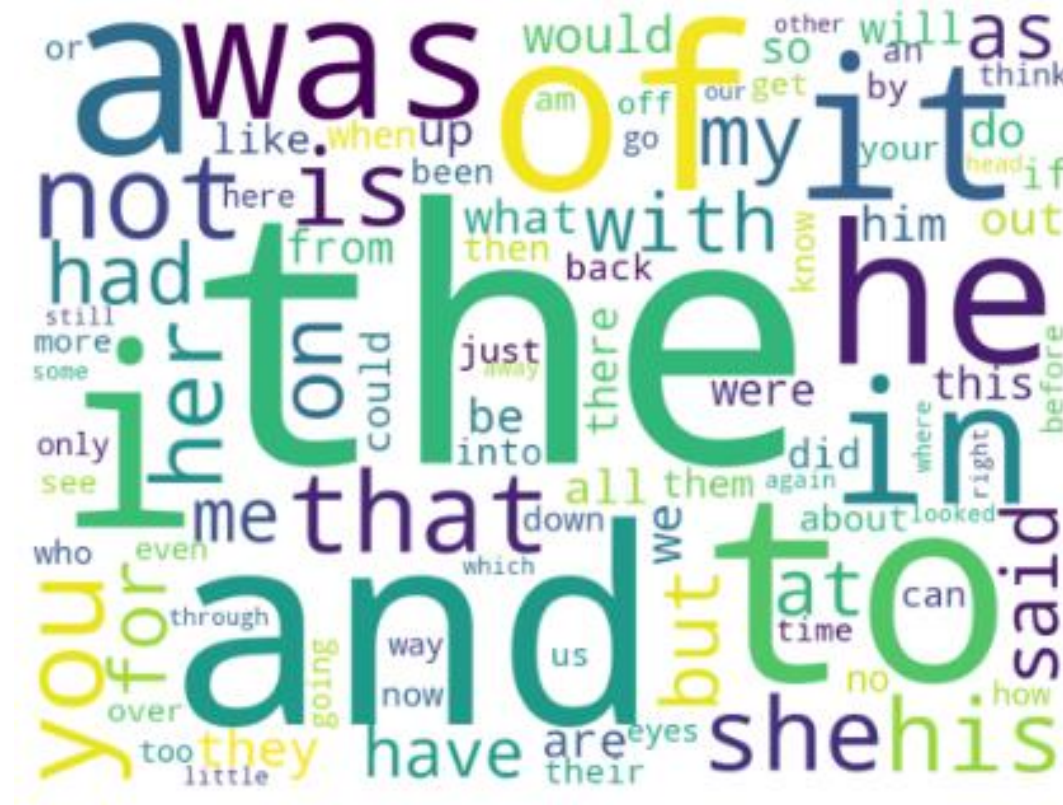


**Books offer a unique opportunity to build vocabulary...**  
**... but they also present a huge challenge!**



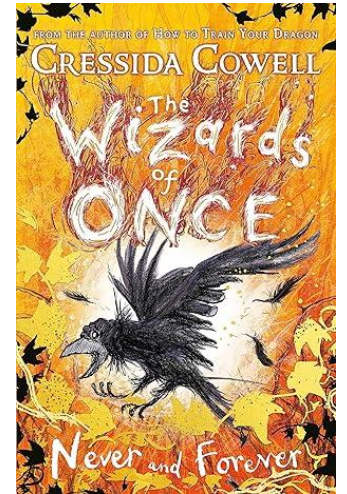
# Few words are used repeatedly

The most common 100 words make up about **54%** of the text



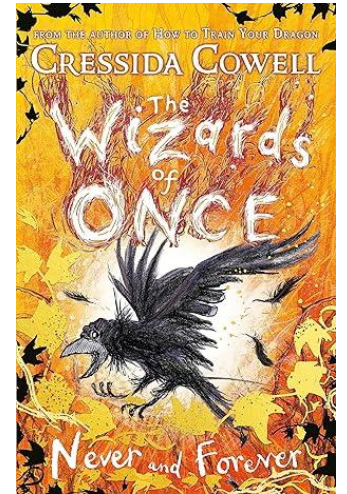
# Few words are used repeatedly

“She glared imperiously through the eyeglass at the entire assembly of assorted creatures and humans, giving them all her hardest stare, and through the glass of the monocle the stare actually stung those it landed on, as if they were being pinched by fairies or bitten by midges, in order to make sure they were truly paying attention”.



# Few words are used repeatedly

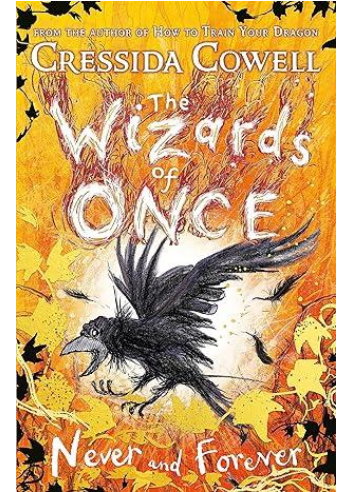
"She through the  
at the of  
and , them all her  
, and through the of the  
the those it  
on, as if they were being by  
or by , in to  
they were ".





# Few words are used repeatedly

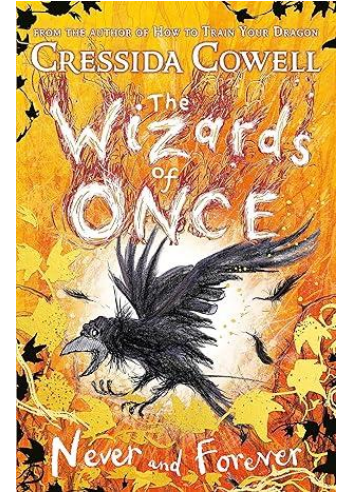
"She through the  
at the of  
and , them all her  
, and through the of the  
the those it  
on, as if they were being by  
or by , in to  
they were ".



- Knowing every second word is not enough to read for meaning

# Few words are used repeatedly

“She glared imperiously through the eyeglass at the entire assembly of assorted creatures and humans, giving them all her hardest stare, and through the glass of the monocle the stare actually stung those it landed on, as if they were being pinched by fairies or bitten by midges, in order to make sure they were truly paying attention”.



- **Knowing every second word is not enough to read for meaning**
- **Important to read widely**

# Half of distinct words have **complex structure**





# Half of distinct words have **complex structure**



# Half of distinct words have **complex structure**



soundlessly



# Half of distinct words have **complex structure**



soundlessly



# Half of distinct words have **complex structure**



soundlessly

# Half of distinct words have **complex structure**



soundlessly



# Half of distinct words have **complex structure**



soundlessly

morphemes

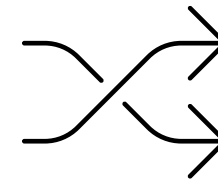


## Morpheme knowledge makes reading faster and more efficient





# Simulating a child reader



useful

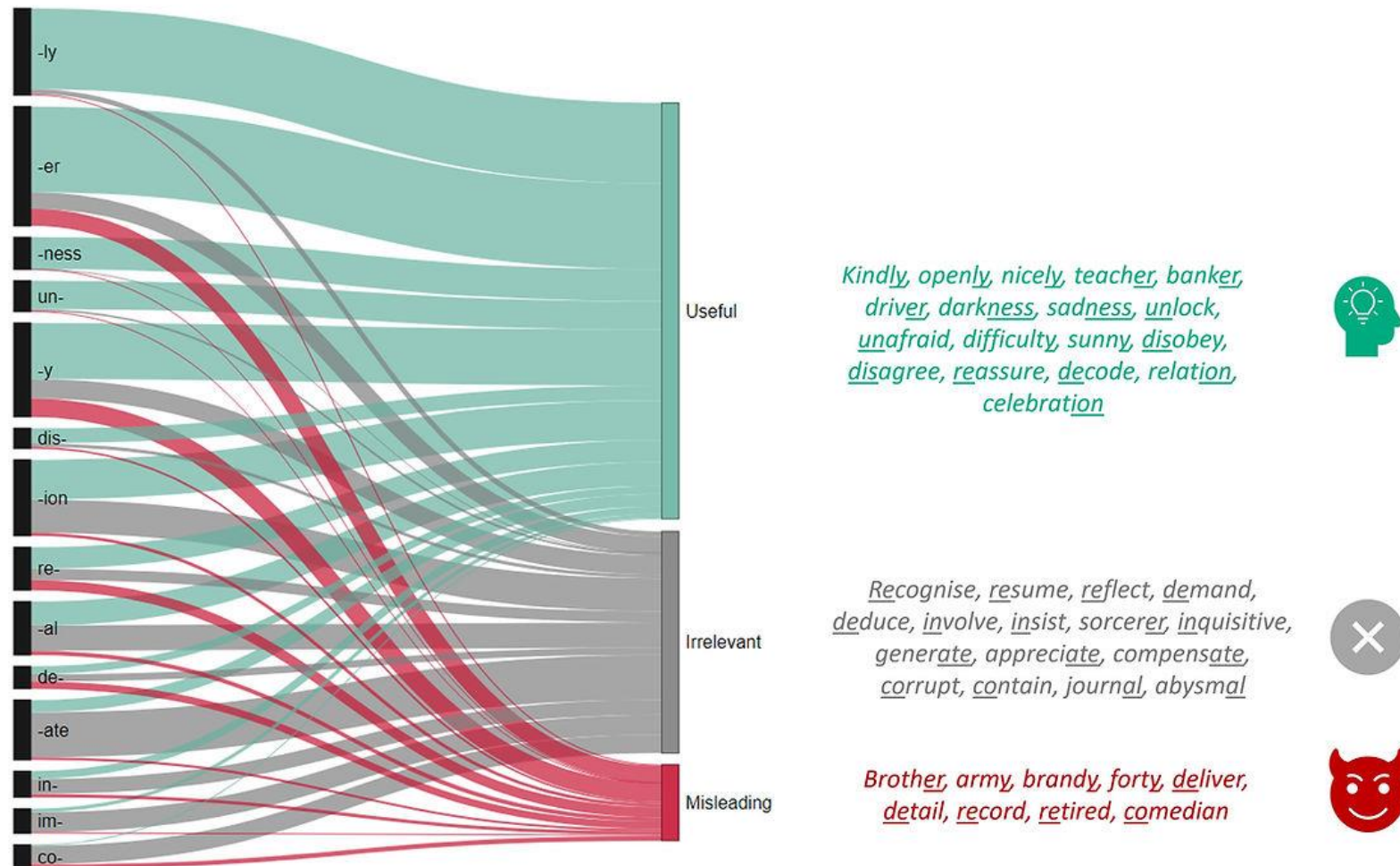


irrelevant

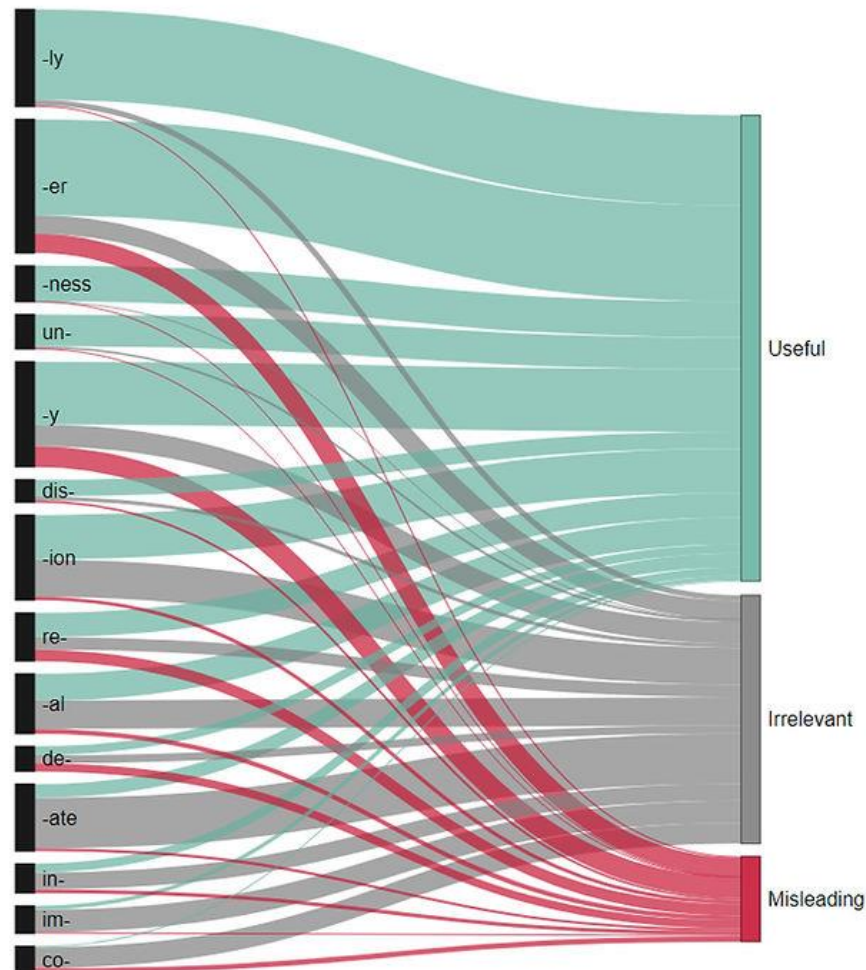


misleading

# Words **do not contribute equally** to learning



# Some words are **useful** for learning morphemes



*Kindly, openly, nicely, teacher, banker,  
driver, darkness, sadness, unlock,  
unafraid, difficulty, sunny, disobey,  
disagree, reassure, decode, relation,  
celebration*



kind + **ly**  
dark + **ness**

*Recognise, resume, reflect, demand,  
deduce, involve, insist, sorcerer, inquisitive,  
generate, appreciate, compensate,  
corrupt, contain, journal, abysmal*

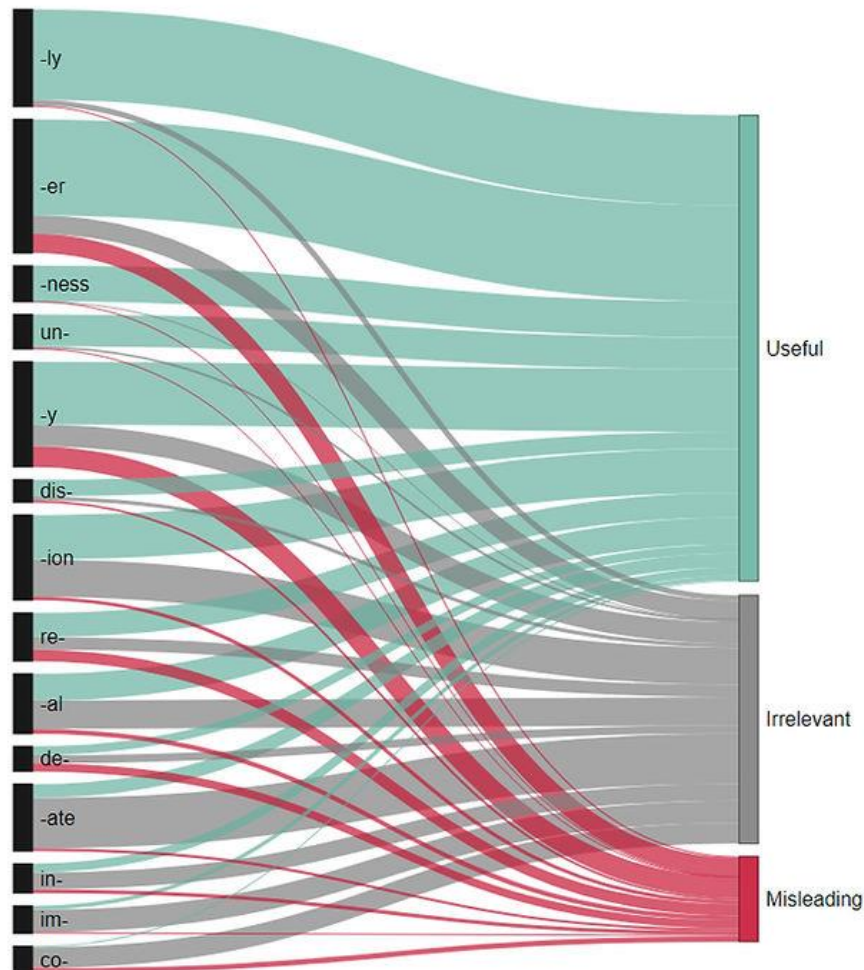


*Brother, army, brandy, forty, deliver,  
detail, record, retired, comedian*





# In many words, morphemes **can't be identified**



*Kindly, openly, nicely, teacher, banker,  
driver, darkness, sadness, unlock,  
unafraid, difficulty, sunny, disobey,  
disagree, reassure, decode, relation,  
celebration*



*Recognise, resume, reflect, demand,  
deduce, involve, insist, sorcerer, inquisitive,  
generate, appreciate, compensate,  
corrupt, contain, journal, abysmal*



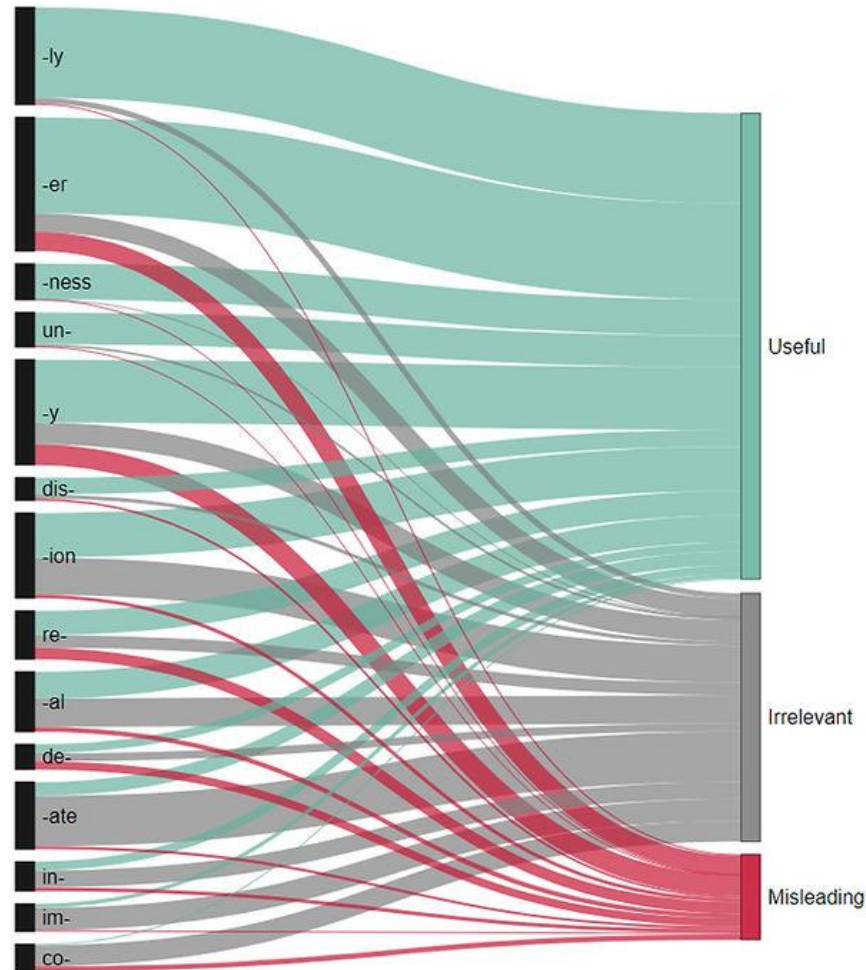
*Brother, army, brandy, forty, deliver,  
detail, record, retired, comedian*



resume  
demand



# Some words only **appear** complex



*Kindly, openly, nicely, teacher, banker,  
driver, darkness, sadness, unlock,  
unafraid, difficulty, sunny, disobey,  
disagree, reassure, decode, relation,  
celebration*



*Recognise, resume, reflect, demand,  
deduce, involve, insist, sorcerer, inquisitive,  
generate, appreciate, compensate,  
corrupt, contain, journal, abysmal*

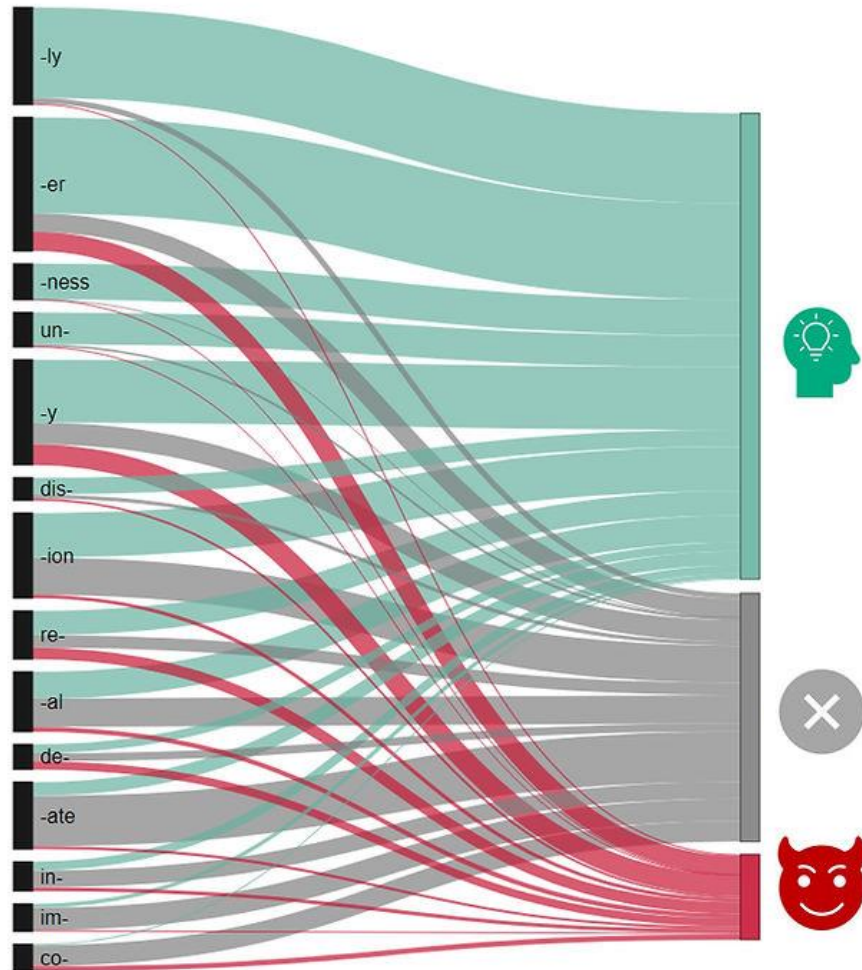


*Brother, army, brandy, forty, deliver,  
detail, record, retired, comedian*



brother  
deliver

# Learning morphemes is **challenging**



Reading alone is unlikely to be enough to learn the meanings of many individual morphemes



Time to rethink morphology instruction?

# Where does this leave us?

- 💡 Books offer a wonderful **opportunity** to build vocabulary
- 💡 Yet, book vocabulary is **challenging** from the get-go
- 💡 Children **need strong phonics skills** to access popular books
- 💡 **Plenty of morphology**, but likely **not enough** to support morpheme **learning**
- 💡 Children need to **read widely** to build reading proficiency
- 💡 **Virtuous cycle** between reading skills and reading motivation



# Further reading

Quarterly Journal of Experimental Psychology  
Volume 77, Issue 12, December 2024, Pages 2418-2438  
© Experimental Psychology Society 2024, [Article Reuse Guidelines](#)  
<https://doi.org/10.1177/17470218241229694>

Sage Journals

Original Article



## 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

Maria Korochkina <sup>1</sup>, Marco Marelli <sup>2</sup>, Marc Brysbaert <sup>3</sup>, and Kathleen Rastle <sup>1</sup>

npj | science of learning

Published in partnership with The University of Queensland

Article



<https://doi.org/10.1038/s41539-025-00313-6>

## Morphology in children's books, and what it means for learning

Check for updates

Maria Korochkina & Kathleen Rastle



Maria Korochkina & Kathy Rastle  
Mar 12, 2024 · 5 min read



## 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...

<https://www.rastlelab.com/blog>



Maria Korochkina & Kathy Rastle  
May 19 · 5 min read



## What can children learn about morphology from reading for fun?

<https://cyp-lex.rastlelab.com/>

CYP-LEX

### Discover what words children and young people encounter when they read

The Children and Young People's Books Lexicon (CYP-LEX) is a database that captures every word used in books popular with children and young people in the UK. It provides insights into how over 100,000 distinct words are used across 1,200 books for ages 7-16. You can use the menu below to look up specific words or pull out words with certain characteristics. Please refer to the READ ME tab for more detail on what information this database provides.

GET STARTED



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*, 77(12), 2415-2438 <https://doi.org/10.1177/17470218241229694>

# Further reading

Quarterly Journal of Experimental Psychology  
Volume 77, Issue 12, December 2024, Pages 2418-2438  
© Experimental Psychology Society 2024, [Article Reuse Guidelines](#)  
<https://doi.org/10.1177/17470218241229694>

Sage Journals

Original Article



## 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

Maria Korochkina <sup>1</sup>, Marco Marelli <sup>2</sup>, Marc Brysbaert <sup>3</sup>, and Kathleen Rastle <sup>1</sup>

npj | science of learning

Published in partnership with The University of Queensland

Article



<https://doi.org/10.1038/s41539-025-00313-6>

## Morphology in children's books, and what it means for learning

Check for updates

Maria Korochkina & Kathleen Rastle



Maria Korochkina & Kathy Rastle  
Mar 12, 2024 · 5 min read



## 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...

<https://www.rastlelab.com/blog>



Maria Korochkina & Kathy Rastle  
May 19 · 5 min read



## What can children learn about morphology from reading for fun?

<https://cyp-lex.rastlelab.com/>

CYP-LEX

### Discover what words children and young people encounter when they read

The Children and Young People's Books Lexicon (CYP-LEX) is a database that captures every word used in books popular with children and young people in the UK. It provides insights into how over 100,000 distinct words are used across 1,200 books for ages 7-16. You can use the menu below to look up specific words or pull out words with certain characteristics. Please refer to the READ ME tab for more detail on what information this database provides.

GET STARTED



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*, 77(12), 2415-2438 <https://doi.org/10.1177/17470218241229694>

# Further reading

Quarterly Journal of Experimental Psychology  
Volume 77, Issue 12, December 2024, Pages 2418-2438  
© Experimental Psychology Society 2024, [Article Reuse Guidelines](#)  
<https://doi.org/10.1177/17470218241229694>

Sage Journals

Original Article



## 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

Maria Korochkina <sup>1</sup>, Marco Marelli <sup>2</sup>, Marc Brysbaert <sup>3</sup>, and Kathleen Rastle <sup>1</sup>

npj | science of learning

Published in partnership with The University of Queensland

Article



<https://doi.org/10.1038/s41539-025-00313-6>

## Morphology in children's books, and what it means for learning

Check for updates

Maria Korochkina & Kathleen Rastle



Maria Korochkina & Kathy Rastle  
Mar 12, 2024 · 5 min read



## 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...

<https://www.rastlelab.com/blog>



Maria Korochkina & Kathy Rastle  
May 19 · 5 min read



## What can children learn about morphology from reading for fun?

<https://cyp-lex.rastlelab.com/>

CYP-LEX

### Discover what words children and young people encounter when they read

The Children and Young People's Books Lexicon (CYP-LEX) is a database that captures every word used in books popular with children and young people in the UK. It provides insights into how over 100,000 distinct words are used across 1,200 books for ages 7-16. You can use the menu below to look up specific words or pull out words with certain characteristics. Please refer to the READ ME tab for more detail on what information this database provides.

GET STARTED



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*, 77(12), 2415-2438 <https://doi.org/10.1177/17470218241229694>



# Further reading

Quarterly Journal of Experimental Psychology  
Volume 77, Issue 12, December 2024, Pages 2418-2438  
© Experimental Psychology Society 2024, [Article Reuse Guidelines](#)  
<https://doi.org/10.1177/17470218241229694>

Sage Journals

Original Article



## 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

Maria Korochkina <sup>1</sup>, Marco Marelli <sup>2</sup>, Marc Brysbaert <sup>3</sup>, and Kathleen Rastle <sup>1</sup>

npj | science of learning

Published in partnership with The University of Queensland

Article



<https://doi.org/10.1038/s41539-025-00313-6>

## Morphology in children's books, and what it means for learning

Check for updates

Maria Korochkina & Kathleen Rastle



Maria Korochkina & Kathy Rastle  
Mar 12, 2024 · 5 min read



## 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...

<https://www.rastlelab.com/blog>



Maria Korochkina & Kathy Rastle  
May 19 · 5 min read



## What can children learn about morphology from reading for fun?

<https://cyp-lex.rastlelab.com/>

CYP-LEX

### Discover what words children and young people encounter when they read

The Children and Young People's Books Lexicon (CYP-LEX) is a database that captures every word used in books popular with children and young people in the UK. It provides insights into how over 100,000 distinct words are used across 1,200 books for ages 7-16. You can use the menu below to look up specific words or pull out words with certain characteristics. Please refer to the READ ME tab for more detail on what information this database provides.

GET STARTED



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*, 77(12), 2415-2438 <https://doi.org/10.1177/17470218241229694>



# Thank you!

[Maria.Korochkina@rhul.ac.uk](mailto:Maria.Korochkina@rhul.ac.uk)

<https://mariakna.github.io/>



Economic  
and Social  
Research Council

