# What can children learn about English morphology through book reading?

**Maria Korochkina**[1] | **Marco Marelli**[2] | **Marc Brysbaert**[3] | **Kathy Rastle**[1]

[1]Department of Psychology, Royal Holloway, University of London, UK

[2]Department of Psychology, University of Milano-Bicocca, Italy

[3]Department of Experimental Psychology, Ghent University, Belgium


EPS meeting
University of York

5 July 2024

# The power of morphology

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]
- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words. . .

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]
- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words...
  - 'un-' + [*stem*] → 'not [stem]' (e.g., *unclean*, *unfair*)

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]

- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words...
  - 'un-' + [*stem*] → 'not [stem]' (e.g., *unclean*, *unfair*)
  - [stem] + '-ify' → 'make [stem]' (e.g., *simplify*, *purify*)

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]
- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words...
  - 'un-' + [*stem*] → 'not [stem]' (e.g., *unclean*, *unfair*)
  - [stem] + '-ify' → 'make [stem]' (e.g., *simplify*, *purify*)
- ...and for computing the meanings of *unfamiliar* words [3, 4]

# The power of morphology

- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]
- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words...
  - 'un-' + [*stem*] → 'not [stem]' (e.g., *unclean*, *unfair*)
  - [stem] + '-ify' → 'make [stem]' (e.g., *simplify*, *purify*)
- ...and for computing the meanings of *unfamiliar* words [3, 4]
  - 'bright' + '-ify' → *brightify*

# The power of morphology

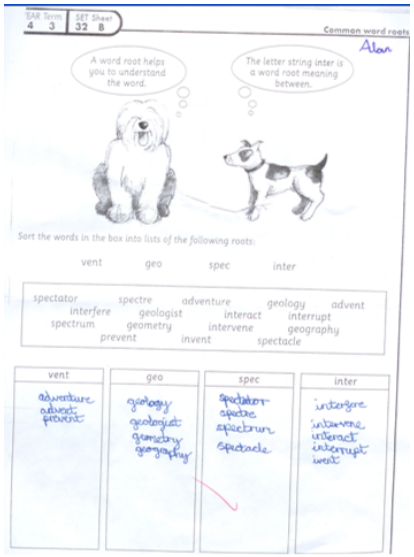
- In languages like English, most words consist of several elements that are themselves meaningful — *morphemes* [1, 2]
- Morpheme knowledge is crucial for quick access to the meanings of *known* printed words...
  - 'un-' + [*stem*] → 'not [stem]' (e.g., *unclean*, *unfair*)
  - [stem] + '-ify' → 'make [stem]' (e.g., *simplify*, *purify*)
- ...and for computing the meanings of *unfamiliar* words [3, 4]
  - 'bright' + '-ify' → *brightify*
- Yet, readers do not show evidence of morpheme knowledge in online reading tasks until mid-to-late adolescence [5, 6]

# Morpheme knowledge & reading

# Morpheme knowledge & reading

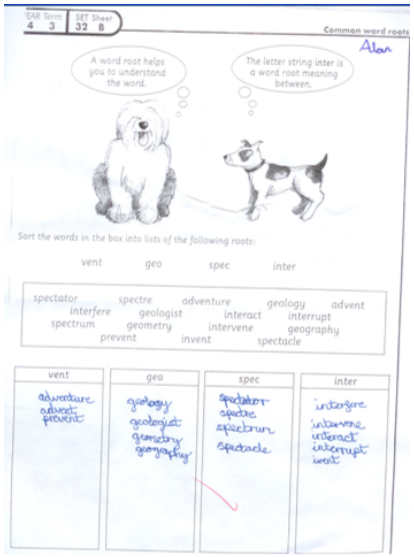
- Little morphological instruction in school

- Little morphological instruction in school
  - Teacher knowledge often sparse & patchy [7]

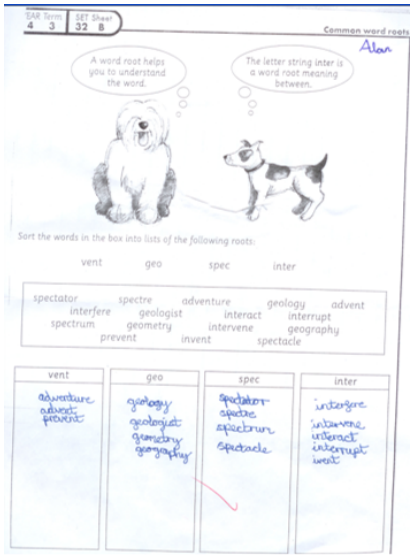
# Morpheme knowledge & reading

- Little morphological instruction in school
  - Teacher knowledge often sparse & patchy [7]

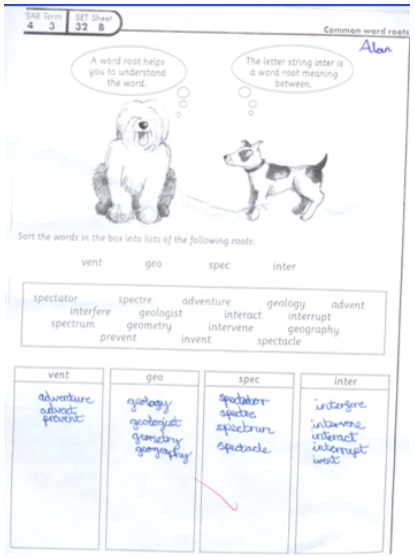

- Form–meaning relationship more salient in written language

- Little morphological instruction in school
  - Teacher knowledge often sparse & patchy [7]


- Form–meaning relationship more salient in written language
  - *bonus*, *atlas*, *service*, *princess* vs. *harazdous* [8–11]

- Little morphological instruction in school
  - Teacher knowledge often sparse & patchy [7]


- Form–meaning relationship more salient in written language
  - *bonus*, *atlas*, *service*, *princess* vs. *harazdous* [8–11]

→ Morpheme knowledge has to be acquired via **text experience**

# Pre-requisites for morpheme learning

| | |
|---|---|
| unknown | arrival |
| unfair | bridal |
| unafraid | principal |
| unlikely | electoral |
| unfix | burial |
| unselfish | arrival |
| unconvinced | withdrawal |

# Pre-requisites for morpheme learning

| | |
|---|---|
| unknown | arrival |
| unfair | bridal |
| unafraid | principal |
| unlikely | electoral |
| unfix | burial |
| unselfish | arrival |
| unconvinced | withdrawal |

- Must have consistent meaning transformation [12]

# Pre-requisites for morpheme learning

| | |
|---|---|
| unknown | arrival |
| unfair | bridal |
| unafraid | principal |
| unlikely | electoral |
| unfix | burial |
| unselfish | arrival |
| unconvinced | withdrawal |

- Must have consistent meaning transformation [12]
- Must occur with a high number of distinct stems [12]

# Pre-requisites for morpheme learning

| | |
|---|---|
| unknown | arrival |
| unfair | bridal |
| unafraid | principal |
| unlikely | electoral |
| unfix | burial |
| unselfish | arrival |
| unconvinced | withdrawal |

- Must have consistent meaning transformation [12]
- Must occur with a high number of distinct stems [12]

But what do children experience **in the wild**?

# The CYP-LEX database

## 1,200 books popular with British children & young people

# The CYP-LEX database

1,200 books popular with British children & young people

7–9


10–12


13+

# The CYP-LEX database

## 1,200 books popular with British children & young people

7–9


10–12


13+


- 400 books per age band

# The CYP-LEX database

1,200 books popular with British children & young people

7–9


10–12


13+


- 400 books per age band
- 70+ mln tokens; 105,694 types

# The CYP-LEX database

1,200 books popular with British children & young people

7–9


10–12


13+


- 400 books per age band
- 70+ mln tokens; 105,694 types
- See Korochkina et al., 2024, *QJEP*, for more detail & link [13]

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

- Roughly half of word types are complex

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

- Roughly half of word types are complex
- % complex words increases with book target age

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

- Roughly half of word types are complex
- % complex words increases with book target age

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

- Roughly half of word types are complex
- % complex words increases with book target age

# Many complex words in books

Based on MorphoLex (68,624 unique words) [14]

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* distinct words | 52,851 | 70,945 | 90,980 |
| *N* words in MorphoLex | 39,149 | 47,363 | 54,557 |
| *N* (%) complex words | 17,634 (**45%**) | 22,564 (**48%**) | 27,554 (**51%**) |
| *N* (%) prefixed words | 4,775 (**27%**) | 6,328 (**28%**) | 8,105 (**29%**) |
| *N* (%) suffixed words | 11,559 (**66%**) | 14,865 (**66%**) | 18,587 (**67%**) |

- Roughly half of word types are complex
- % complex words increases with book target age
- More exposure to suffixed than to prefixed words

# Most 'new' words in books are complex

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]
- CBBC: aimed at children aged 6–12; 58,691 types

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]
- CBBC: aimed at children aged 6–12; 58,691 types

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* words missing from CBBC | 8,280 | 14,050 | 20,105 |
| *N* (%) complex words | 4,924 (**59%**) | 8,562 (**61%**) | 12,894 (**64%**) |
| *N* words missing from SUBTLEX-UK | 1,211 | 2,450 | 4,602 |
| *N* (%) complex words | 888 (**73%**) | 1,796 (**73%**) | 3,514 (**76%**) |

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]
- CBBC: aimed at children aged 6–12; 58,691 types

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* words missing from CBBC | 8,280 | 14,050 | 20,105 |
| *N* (%) complex words | 4,924 (**59%**) | 8,562 (**61%**) | 12,894 (**64%**) |
| *N* words missing from SUBTLEX-UK | 1,211 | 2,450 | 4,602 |
| *N* (%) complex words | 888 (**73%**) | 1,796 (**73%**) | 3,514 (**76%**) |

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]
- CBBC: aimed at children aged 6–12; 58,691 types

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* words missing from CBBC | 8,280 | 14,050 | 20,105 |
| *N* (%) complex words | 4,924 (**59%**) | 8,562 (**61%**) | 12,894 (**64%**) |
| *N* words missing from SUBTLEX-UK | 1,211 | 2,450 | 4,602 |
| *N* (%) complex words | 888 (**73%**) | 1,796 (**73%**) | 3,514 (**76%**) |

- Most words not encountered on TV are complex

# Most 'new' words in books are complex

- SUBTLEX-UK: 9 channels over 3 years; 159,235 types [15]
- CBBC: aimed at children aged 6–12; 58,691 types

| | 7–9 | 10–12 | 13+ |
|---|---|---|---|
| *N* words missing from CBBC | 8,280 | 14,050 | 20,105 |
| *N* (%) complex words | 4,924 (**59%**) | 8,562 (**61%**) | 12,894 (**64%**) |
| *N* words missing from SUBTLEX-UK | 1,211 | 2,450 | 4,602 |
| *N* (%) complex words | 888 (**73%**) | 1,796 (**73%**) | 3,514 (**76%**) |

- Most words not encountered on TV are complex
- Books are an important source of morpheme information

# Few complex words used repeatedly

# Few complex words used repeatedly

| | | 7–9 | 10–12 | 13+ |
|---|---|---|---|---|
| Complex | Frequency $\leq 5$ | 50% | 42% | 35% |
| | Frequency $\geq 100$ | 8% | 11% | 15% |
| Monomorphemic | Frequency $\leq 5$ | 30% | 24% | 18% |
| | Frequency $\geq 100$ | 23% | 29% | 36% |

# Few complex words used repeatedly

| | | 7–9 | 10–12 | 13+ |
|---|---|---|---|---|
| Complex | Frequency $\leq 5$ | 50% | 42% | 35% |
| | Frequency $\geq 100$ | 8% | 11% | 15% |
| Monomorphemic | Frequency $\leq 5$ | 30% | 24% | 18% |
| | Frequency $\geq 100$ | 23% | 29% | 36% |

# Few complex words used repeatedly

| | | 7–9 | 10–12 | 13+ |
|---|---|---|---|---|
| Complex | Frequency $\leq 5$ | 50% | 42% | 35% |
| | Frequency $\geq 100$ | 8% | 11% | 15% |
| Monomorphemic | Frequency $\leq 5$ | 30% | 24% | 18% |
| | Frequency $\geq 100$ | 23% | 29% | 36% |

# Few complex words used repeatedly

| | | 7–9 | 10–12 | 13+ |
|---|---|---|---|---|
| Complex | Frequency $\leq 5$ | 50% | 42% | 35% |
| | Frequency $\geq 100$ | 8% | 11% | 15% |
| Monomorphemic | Frequency $\leq 5$ | 30% | 24% | 18% |
| | Frequency $\geq 100$ | 23% | 29% | 36% |

- Morphologically simple words are much more common than complex words

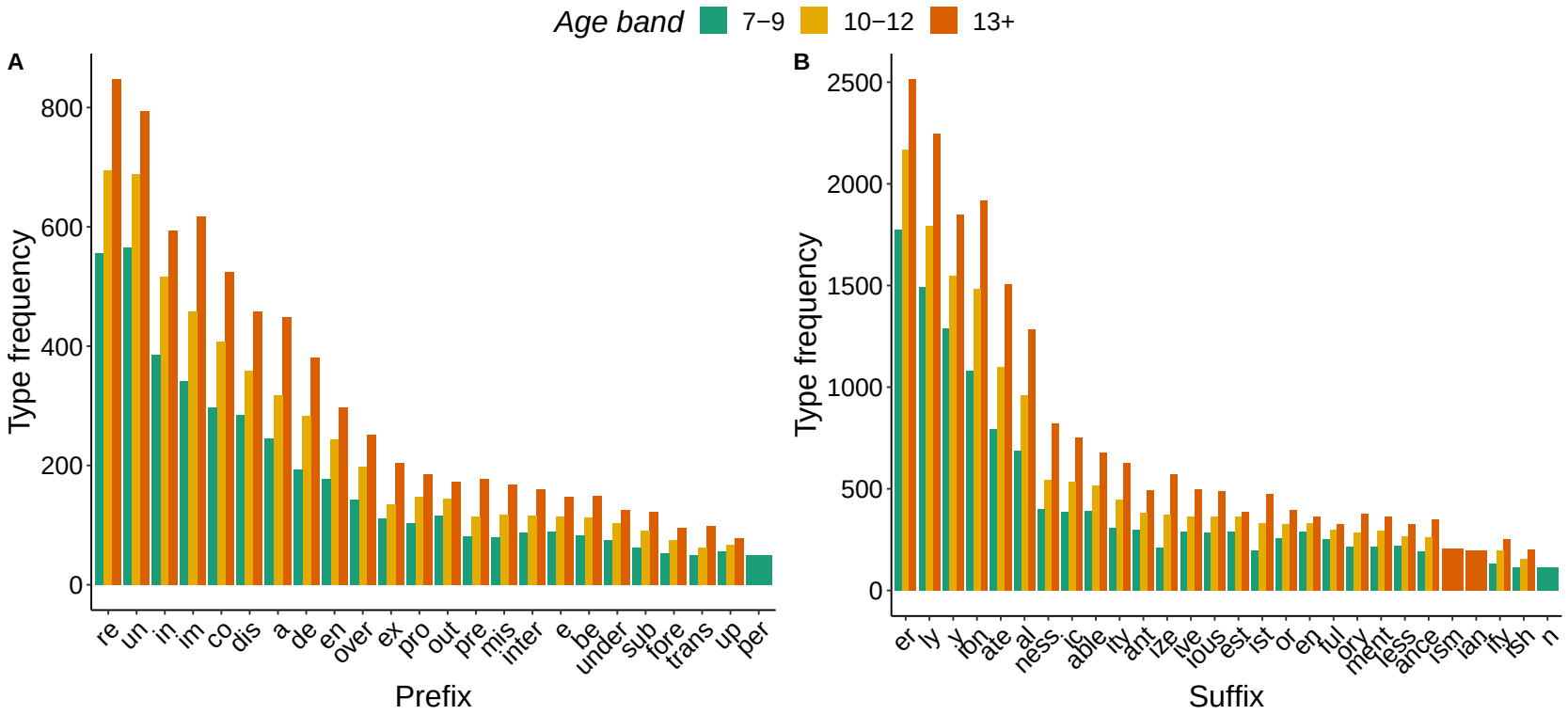
# Few complex words used repeatedly

| | | 7–9 | 10–12 | 13+ |
|---|---|---|---|---|
| Complex | Frequency $\leq 5$ | 50% | 42% | 35% |
| | Frequency $\geq 100$ | 8% | 11% | 15% |
| Monomorphemic | Frequency $\leq 5$ | 30% | 24% | 18% |
| | Frequency $\geq 100$ | 23% | 29% | 36% |

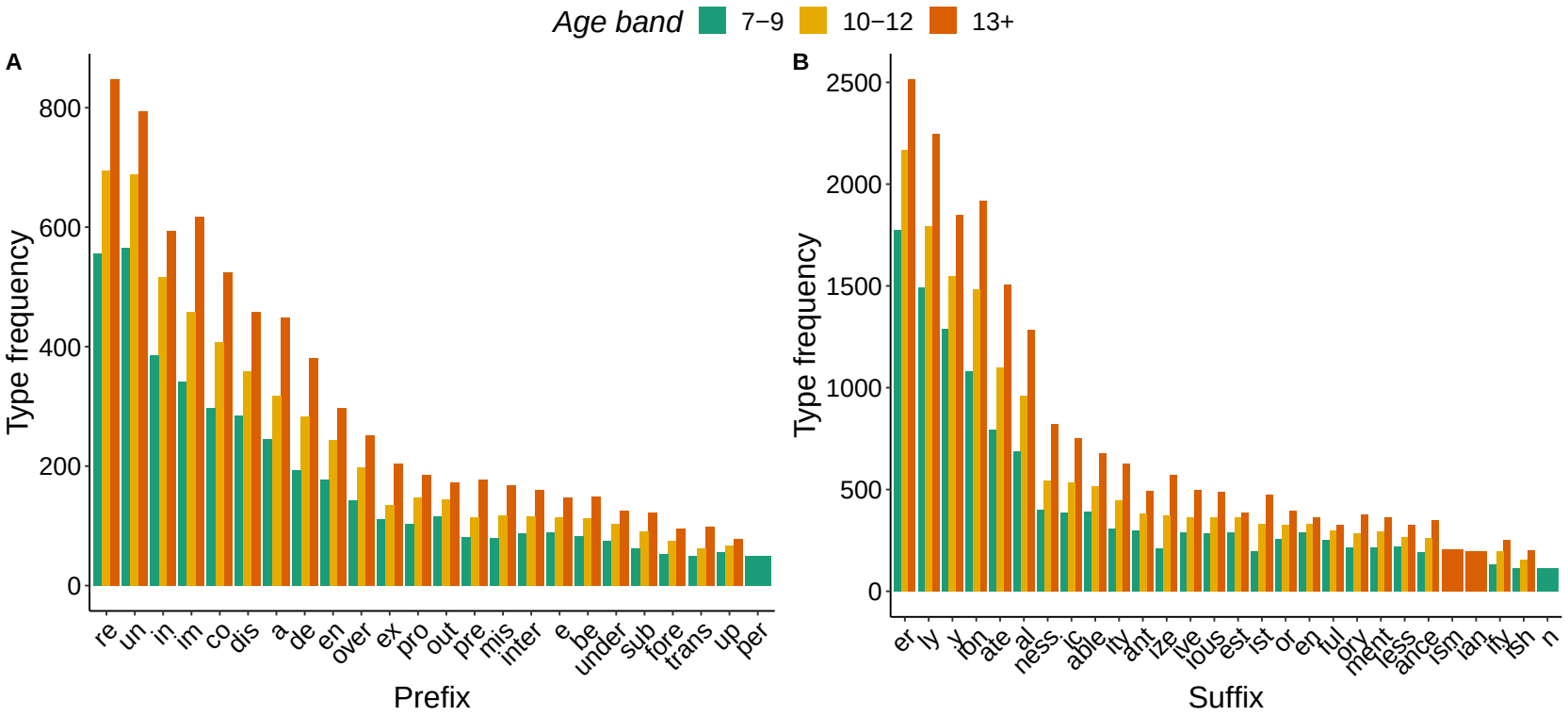
- Morphologically simple words are much more common than complex words
- Readers encounter many morphologically complex words, but few are repeated frequently

# Only a few affixes encountered often

# Only a few affixes encountered often

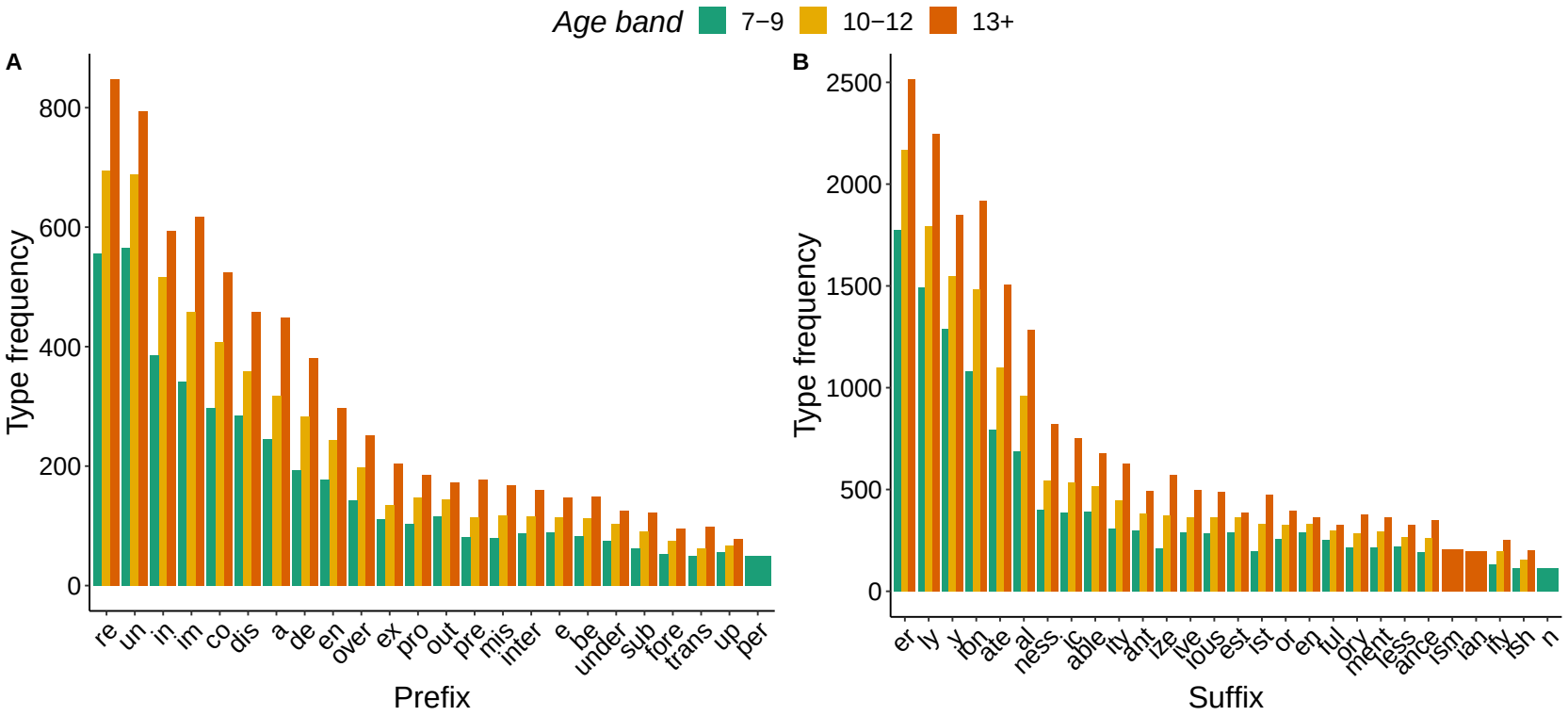


# Only a few affixes encountered often



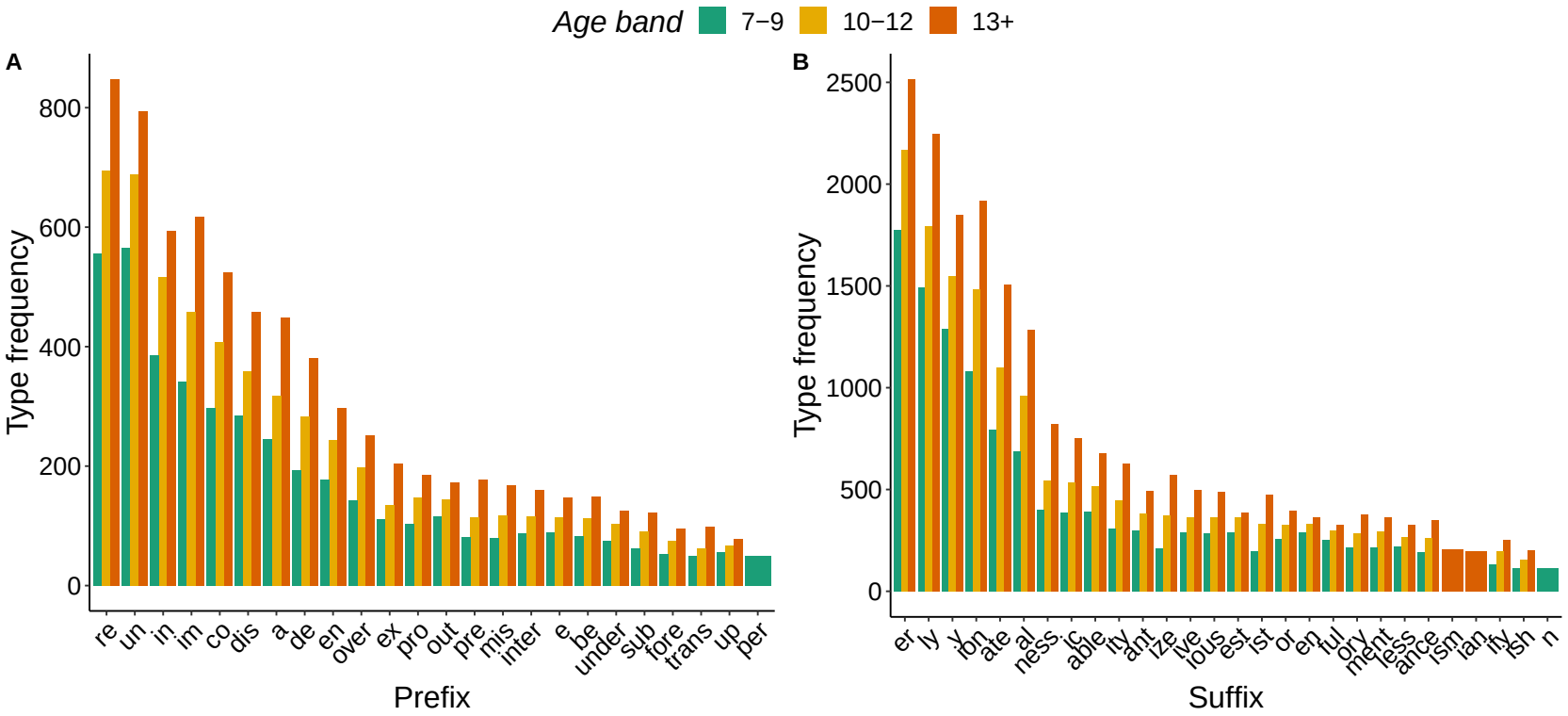

- Prefixes: *re-*, *un-*, *in-*, *co-*

# Only a few affixes encountered often




- Prefixes: *re-*, *un-*, *in-*, *co-*
- Suffixes: *-er*, *-ly*, *-y*, *-ion*, *-ate*, *-al*, *-ness*, *-able*, *-ic*

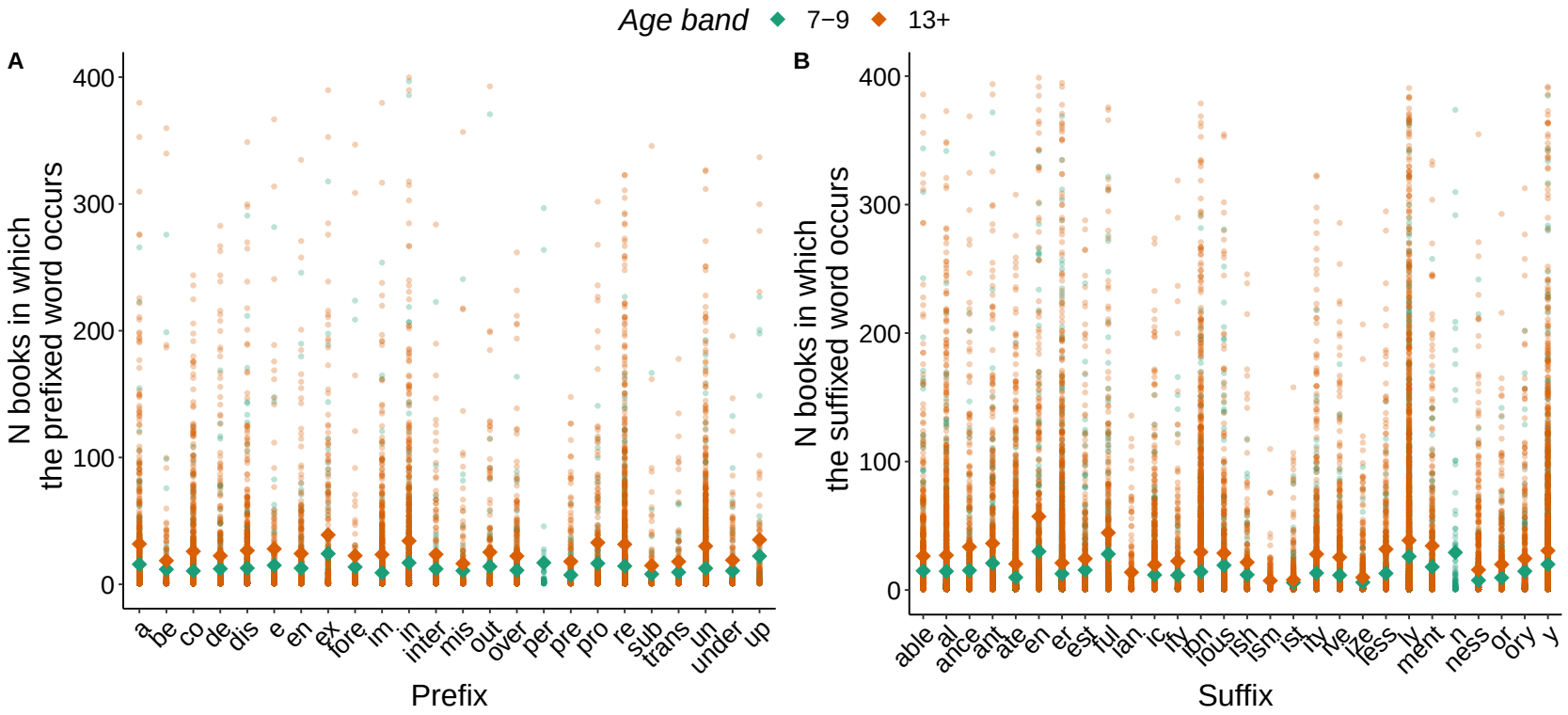
# Only a few affixes encountered often



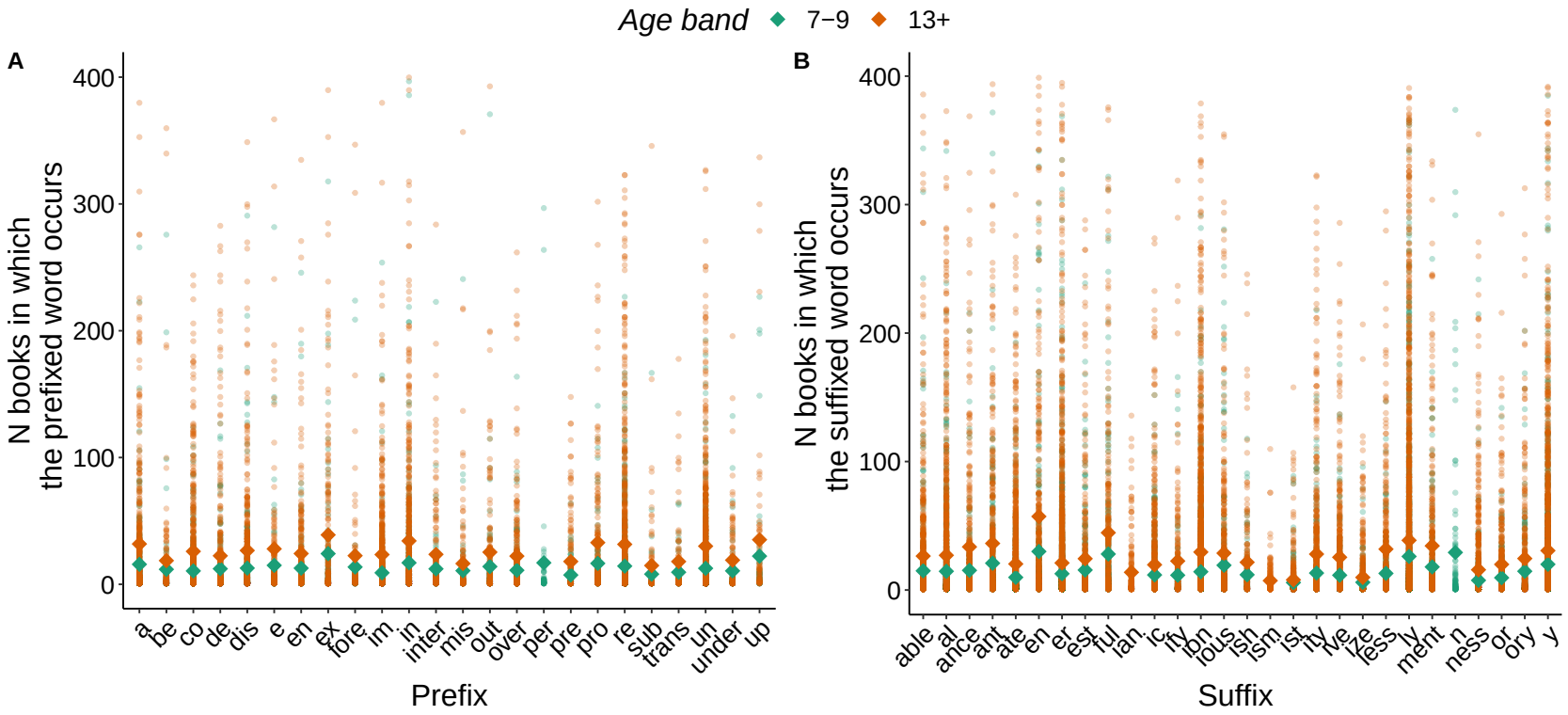

- Prefixes: *re-*, *un-*, *in-*, *co-*
- Suffixes: *-er*, *-ly*, *-y*, *-ion*, *-ate*, *-al*, *-ness*, *-able*, *-ic*
- Limited exposure to multiple types before 13+ texts

# Affixes sparsely represented across books

# Affixes sparsely represented across books

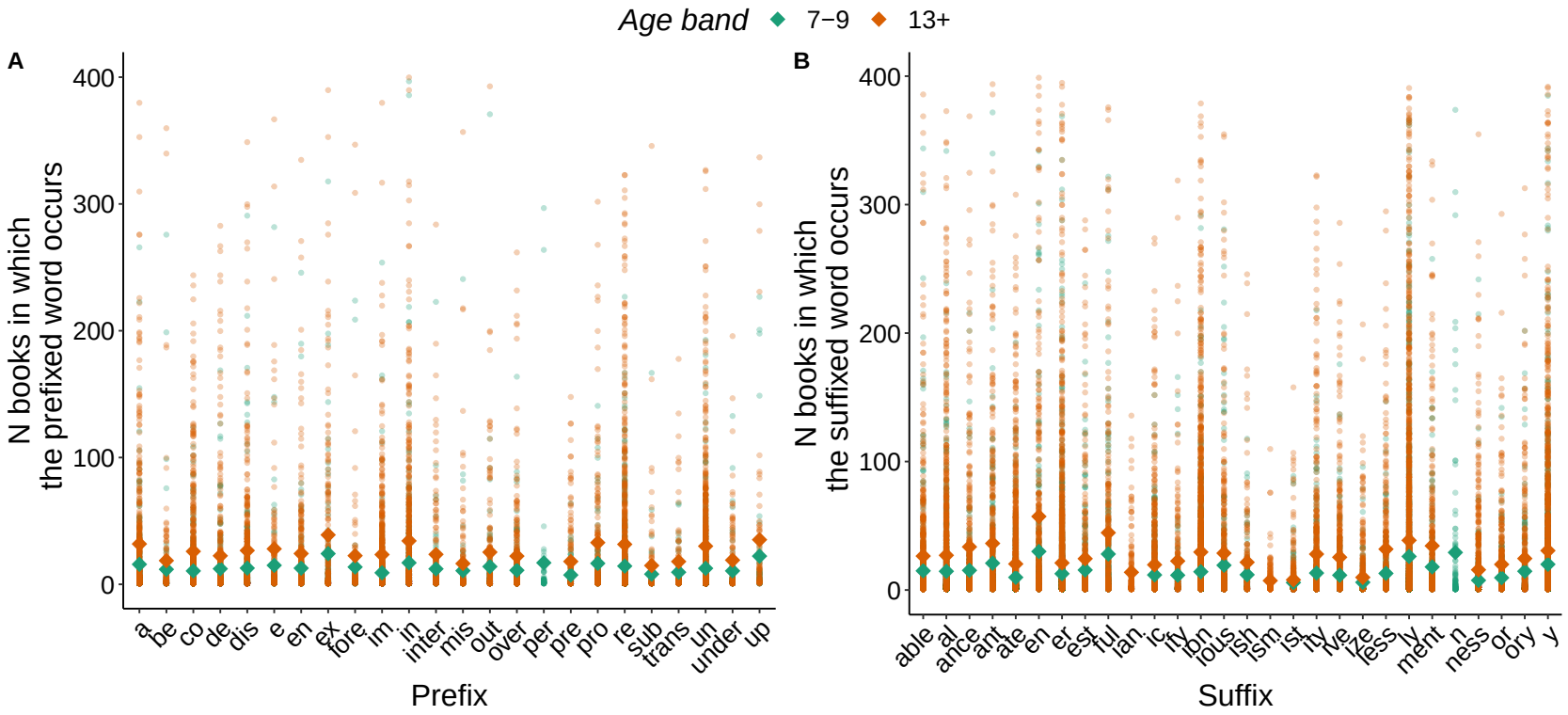


# Affixes sparsely represented across books




- The average affixed word does not occur in many books

# Affixes sparsely represented across books




- The average affixed word does not occur in many books
- Few affixes have reasonable representation before 13+ texts

# Are morphemes easy to 'find'?

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books
- MorphoLex defines morphemes based on etymology

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books
- MorphoLex defines morphemes based on etymology
  - *per*fect, *in*clude, *dis*tribute, *sus*tain

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books
- MorphoLex defines morphemes based on etymology
  - *per*fect, *in*clude, *dis*tribute, *sus*tain
- Yet, children's etymological knowledge is limited

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books
- MorphoLex defines morphemes based on etymology
  - *per*fect, *in*clude, *dis*tribute, *sus*tain
- Yet, children's etymological knowledge is limited
- Exposure to morphology should not be judged on the basis of etymological counts

# Are morphemes easy to 'find'?

- Complex words are low in frequency & poorly distributed across books
- MorphoLex defines morphemes based on etymology
  - *per*fect, *in*clude, *dis*tribute, *sus*tain
- Yet, children's etymological knowledge is limited
- Exposure to morphology should not be judged on the basis of etymological counts

How does the picture change if morphemes are defined **orthographically**?

# The Regular Expression approach

# The Regular Expression approach

A string-searching algorithm to detect words that *appear* complex

# The Regular Expression approach

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations

# The Regular Expression approach

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations

# The Regular Expression approach

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations
  - 'sun' + '-y' $\rightarrow$ *sun***n***y*

# The Regular Expression approach

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations
  - 'sun' + '-y' → *sun***n***y*
  - 'solv**e**' + '-able' → *solvable*

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations
  - 'sun' + '-y' → *sun***n***y*
  - 'solv**e**' + '-able' → *solvable*

| | *N* MorphoLex complex | RegEx hits | RegEx FAs |
|---|---|---|---|
| Prefixed | 8,105 | 3,811 (**47%**) | 1,510 (**3.3%**) |
| Suffixed | 18,587 | 8,801 (**47%**) | 1,735 (**4.8%**) |

A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations
  - 'sun' + '-y' → *sun***n***y*
  - 'solv**e**' + '-able' → *solvable*

| | *N* MorphoLex complex | RegEx hits | RegEx FAs |
|---|---|---|---|
| Prefixed | 8,105 | 3,811 (**47%**) | 1,510 (**3.3%**) |
| Suffixed | 18,587 | 8,801 (**47%**) | 1,735 (**4.8%**) |

- Hits low due to 'missing' stems (e.g., *pessimist*) & complex alterations (e.g., 'sub-' + 'tenere' → *sustain*)

# The Regular Expression approach

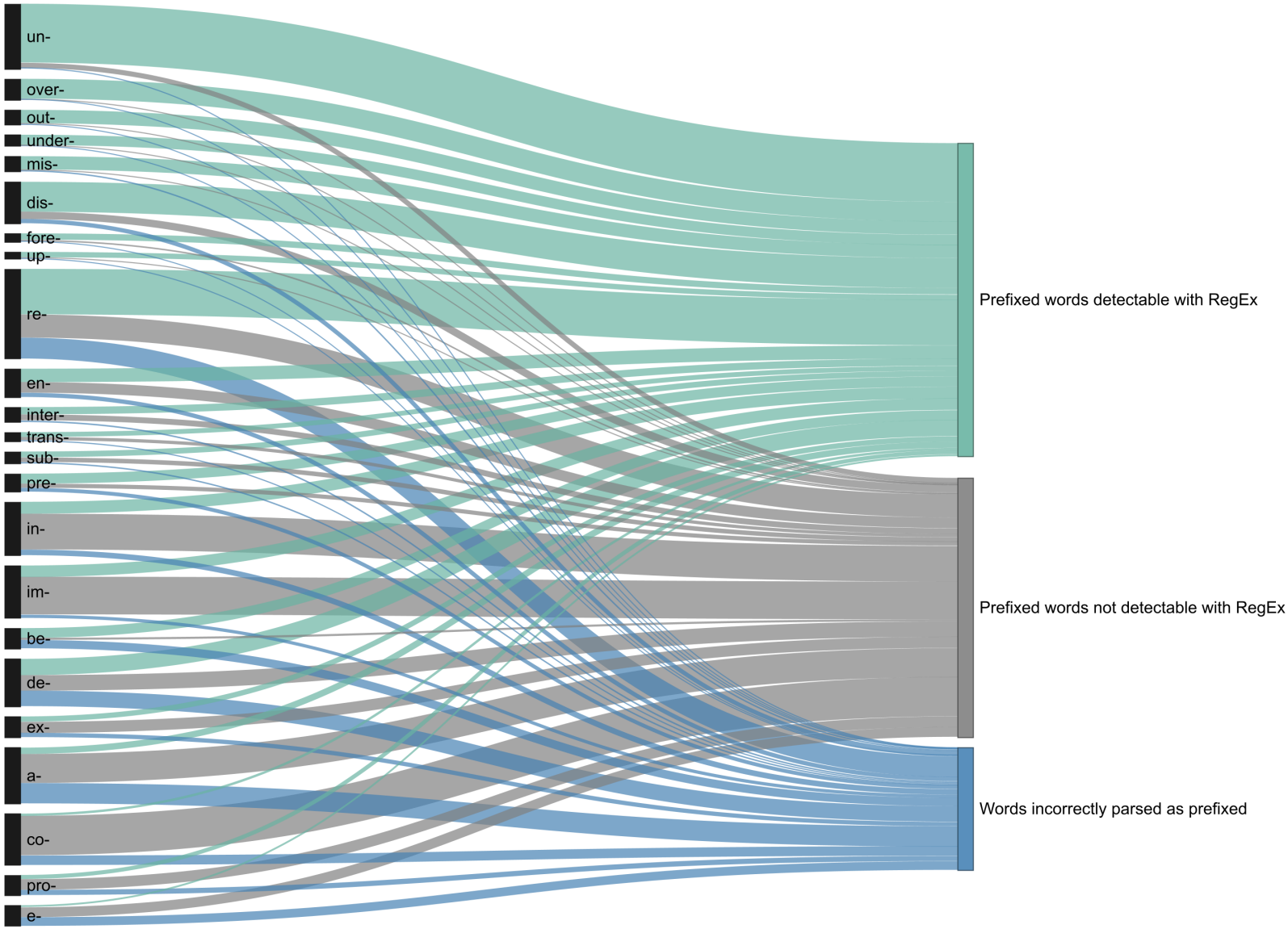
A string-searching algorithm to detect words that *appear* complex

- Recursive search of legal stem–affix combinations
- Sensitive to common orthographic alterations
  - 'sun' + '-y' → *sun***n***y*
  - 'solv**e**' + '-able' → *solvable*

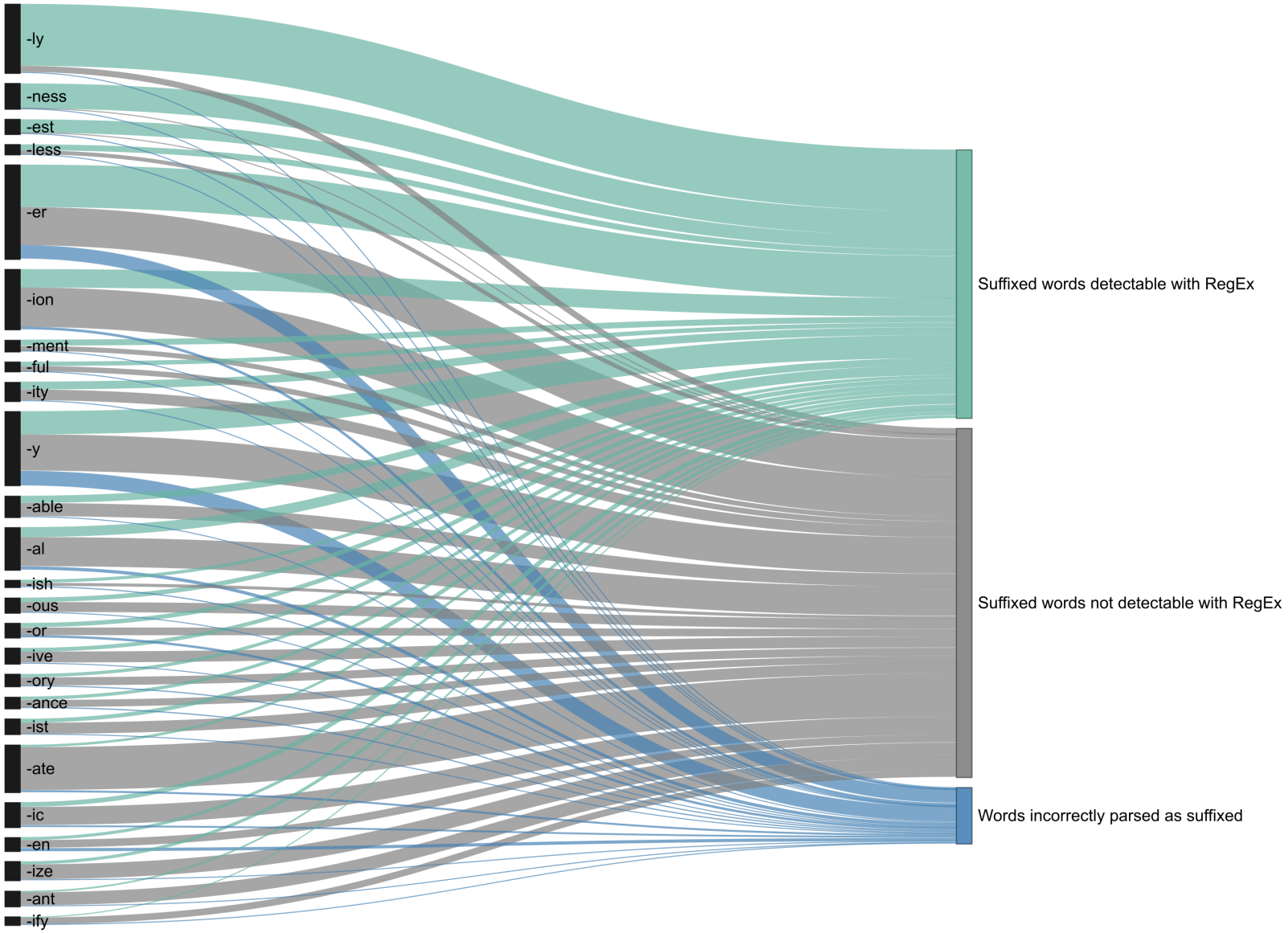
| | *N* MorphoLex complex | RegEx hits | RegEx FAs |
|---|---|---|---|
| Prefixed | 8,105 | 3,811 (**47%**) | 1,510 (**3.3%**) |
| Suffixed | 18,587 | 8,801 (**47%**) | 1,735 (**4.8%**) |

- Hits low due to 'missing' stems (e.g., *pessimist*) & complex alterations (e.g., 'sub-' + 'tenere' → *sustain*)
- False alarms due to pseudoaffixation (e.g., *corner*)

# Some prefixes are much easier to detect than others



# Some suffixes are much easier to detect than others



# Conclusions

# Conclusions

- Complex words comprise a large proportion of words in children's books

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes
  - Frequency, representation across books, detectability

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes
  - Frequency, representation across books, detectability
- Beyond a handful of affixes, morpheme knowledge will be difficult to acquire from text

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes
  - Frequency, representation across books, detectability
- Beyond a handful of affixes, morpheme knowledge will be difficult to acquire from text

$\rightarrow$ Is this the reason why children do not show evidence of morpheme knowledge in online reading tasks until late adolescence?

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes
  - Frequency, representation across books, detectability
- Beyond a handful of affixes, morpheme knowledge will be difficult to acquire from text

$\rightarrow$ Is this the reason why children do not show evidence of morpheme knowledge in online reading tasks until late adolescence?

- Morpheme information is *graded* $\rightarrow$ morpheme knowledge is *graded*

# Conclusions

- Complex words comprise a large proportion of words in children's books
- Substantial variation across affixes
  - Frequency, representation across books, detectability
- Beyond a handful of affixes, morpheme knowledge will be difficult to acquire from text

$\rightarrow$ Is this the reason why children do not show evidence of morpheme knowledge in online reading tasks until late adolescence?

- Morpheme information is *graded* $\rightarrow$ morpheme knowledge is *graded*
- Important to study morpheme properties & their *relationship to learning*

*Thank you!*

[1] P. Angelelli, C. Marinelli, and C. Burani, "The effect of morphology on spelling and reading accuracy: A study on Italian children," *Frontiers in Psychology*, vol. 5, no. Article 1373, pp. 1–10, 2014. DOI: https://doi.org/10.3389/fpsyg.2014.01373.

[2] R. Baayen, R. Piepenbrock, and H. van Rijn, *The CELEX lexical database, (CD-ROM)*. PA: University of Pennsylvania. Linguistic Data Consortium, 1993.

[3] K. Rastle and M. Davis, "Morphological decomposition based on the analysis of orthography," *Language and Cognitive Processes*, vol. 23, no. 7/8, pp. 942–971, 2008. DOI: https://doi.org/10.1080/01690960802069730.

[4] K. Rastle, "The place of morphology in learning to read in English," *Cortex*, vol. 116, pp. 45–54, 2019. DOI: https://doi.org/10.1016/j.cortex.2018.02.008.

[5] N. Dawson, K. Rastle, and J. Ricketts, "Morphological effects in visual word recognition: Children, adolescents, and adults," *Journal of Experimental Psychology: Learning Memory and Cognition*, vol. 44, no. 4, pp. 645–654, 2018. DOI: https://doi.org/10.1037/xlm0000485.

[6] N. Dawson, K. Rastle, and J. Ricketts, "Finding the man amongst many: A developmental perspective on mechanisms of morphological decomposition," *Cognition*, vol. 211, p. 104 605, 2021. DOI: https://doi.org/10.1016/j.cognition.2021.104605.

[7] J. Hurry *et al.*, "Transforming research on morphology into teacher practice," *Research Papers in Education*, vol. 20, pp. 187–206, 2005. DOI: https://doi.org/10.1080/02671520500078291.

[8] K. Berg and M. Aronoff, "Self-organization in the spelling of English suffixes: The emergence of culture out of anarchy," *Language*, vol. 93, no. 1, pp. 37–64, 2017. DOI: https://doi.org/10.1353/lan.2017.0000.

[9] E. Carney, *A survey of English spelling*. London, England: Routledge, 1994.

[10] C. Chomsky, "Reading, writing, and phonology," *Harvard Educational Review*, vol. 40, no. 2, pp. 287–309, 1970. DOI: https://doi.org/10.17763/haer.40.2.y7u0242x76w05624.

[11] K. Rastle, "EPS Mid-Career Prize Lecture 2017: Writing systems, reading, and language," *Quarterly Journal of Experimental Psychology*, vol. 72, no. 4, pp. 677–692, 2019. DOI: https://doi.org/10.1177/1747021819829696.

[12] J. Tamminen, M. Davis, and K. Rastle, "From specific examples to general knowledge in language learning," *Cognitive Psychology*, vol. 79, pp. 1–39, 2015. DOI: https://doi.org/10.1016/j.cogpsych.2015.03.003.

[13] M. Korochkina, M. Marelli, M. Brybaert, and K. Rastle, "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*, 2024. DOI: https://doi.org/10.1177/17470218241229694.

[14] C. H. Sánchez-Gutiérrez, H. Mailhot, S. H. Deacon, and M. A. Wilson, "MorphoLex: A derivational morphological database for 70,000 English words," *Behavior Research Methods*, vol. 50, no. 4, pp. 1568–1580, 2018. DOI: 10.3758/s13428-017-0981-8.

[15] W. J. van Heuven, P. Mandera, E. Keuleers, and M. Brysbaert, "SUBTLEX-UK: A new and improved word frequency database for British English," *Quarterly Journal of Experimental Psychology*, vol. 67, no. 6, pp. 1176–1190, 2014. DOI: 10.1080/17470218.2013.850521.