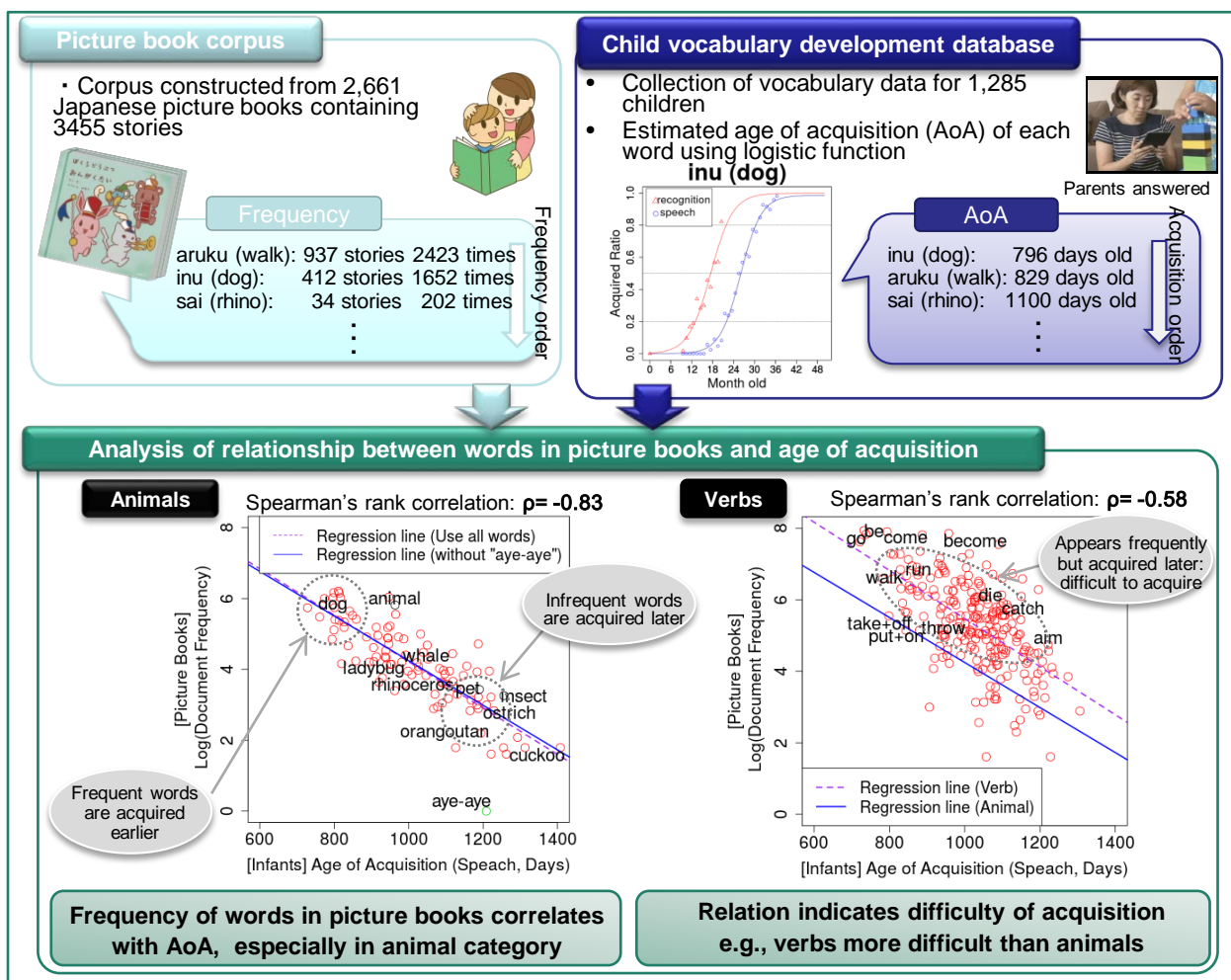# 10 > Early acquired words correlate with picture books
- Relation between picture book corpus and child vocabulary -

Abstract

We constructed a large-scale corpus of Japanese picture books and analyzed the relationship between words that appear in those books and child vocabulary acquisition. Although previous studies have suggested that picture book reading promotes child vocabulary acquisition, the effect has not been fully examined due to lack of picture book corpora. In our study, we found a strong correlation between the age of acquisition for basic-level nouns such as animal names and the frequency of the word in picture books. Verbs required a larger frequency for acquisition compared to basic-level nouns. The relationship between age of acquisition and frequency of words in picture books can indicate the difficulty of acquisition for a given word. Based on these findings, we can improve educational support systems such as recommending picture books according to an individual's vocabulary level.

Picture book corpus

· Corpus constructed from 2,661 Japanese picture books containing 3455 stories

Frequency

aruku (walk): 937 stories 2423 times
inu (dog):   412 stories 1652 times
sai (rhino):   34 stories   202 times

Frequency order

Child vocabulary development database

• Collection of vocabulary data for 1,285 children
• Estimated age of acquisition (AoA) of each word using logistic function

inu (dog)

Parents answered

AoA

inu (dog):   796 days old
aruku (walk): 829 days old
sai (rhino):   1100 days old

Acquisition order

Analysis of relationship between words in picture books and age of acquisition

Animals   Spearman's rank correlation: ρ= -0.83

Infrequent words are acquired later

Frequent words are acquired earlier

Frequency of words in picture books correlates with AoA, especially in animal category

Verbs   Spearman's rank correlation: ρ= -0.58

Appears frequently but acquired later: difficult to acquire

Relation indicates difficulty of acquisition e.g., verbs more difficult than animals

References

[1] S. Fujita, T. Kobayashi, Y. Okumura, T. Hattori, "How do words in picture books affect child vocabulary acquisition? – An analysis of large-scaled corpus in Japanese picture books –," in *Proc. Workshop on Infant Language Development (WILD-2017)*, 2017.

[2] T. Kobayashi, Y. Okumura, Y. Minami, "Collecting Data on Child Vocabulary Development by Vocabulary-checklist Application," in *Proc. IEICE, HCS2015-59, Vol. 115, No. 418, pp.1–6, 2016.*

Contact

Sanae Fujita Linguistic Intelligence, Innovative Communication Laboratory
Email : fujita.sanae(at)lab.ntt.co.jp