Testing a Chunking Account of the Relationship Between Speed of Processing and Children's Vocabulary

Hannah Sawyer[a], Julia Egger[b], Caroline Rowland[c], Julian Pine[a], Christina Bergmann[c], Andrew Jessop[a], Gary Jones[d] & Samantha Durrant[e]

[a]University of Liverpool, [b]University of Vienna, [c]Max Planck Institute for Psycholinguistics, [d]University of Nottingham, [e]University of Manchester

Contact: Hshsawy3@liverpool.ac.uk, Twitter: HSawyer01

Introduction

- The speed with which children process familiar words is robustly associated with vocabulary size (see Peter et al. 2019; for a review).
- There are two possible explanations:

1. Intrinsic processing speed differences.	2. Experience-dependent processing speed differences.
• Some children process familiar words faster. • **This relationship reflects the effects of processing on learning.** • Faster processing of familiar words frees up resources that can be dedicated to learning new words = bigger vocabulary.	• **This relationship reflects the effects of vocabulary on processing.** • Children exposed to more language have larger vocabularies. • Therefore, they are better at processing familiar words as they have a larger bank of linguistic knowledge which they can use to efficiently process and comprehend incoming utterances.

Study 1 Aim: To see whether a computational model (CLASSIC) can simulate the relationship between processing speed and vocabulary with only experience-related differences.

Study 1: CLASSIC

- CLASSIC (Jones et al. 2007, 2014) is a derivative of the EPAM/CHREST architecture (Gobet et al., 2001) that uses an associative chunking mechanism to process incoming utterances and expand the current knowledge base.

Chunking = a process where multiple individual elements are compressed and recoded into a single perceptual unit (Cowan, 2010; Gobet et al., 2001; Miller, 1956), such as when strings of phonemes are grouped to form a word.

- CLASSIC gradually learns the phonological forms of words and word sequences by binding adjacent strings in the input into increasingly larger chunks as exposure to these strings increases.

/c/ /a/ /t/ → /c/ /at/ → /cat/

- CLASSIC was trained on different sized samples of English caregiver speech and was given the same vocabulary and processing speed tasks as in Peter et al. (2019).
- CLASSIC simulated the relationship between processing speed and vocabulary - models with greater exposure to language input had a larger vocabulary and were able to represent information in fewer chunks (= faster processing) (see Figure 1):

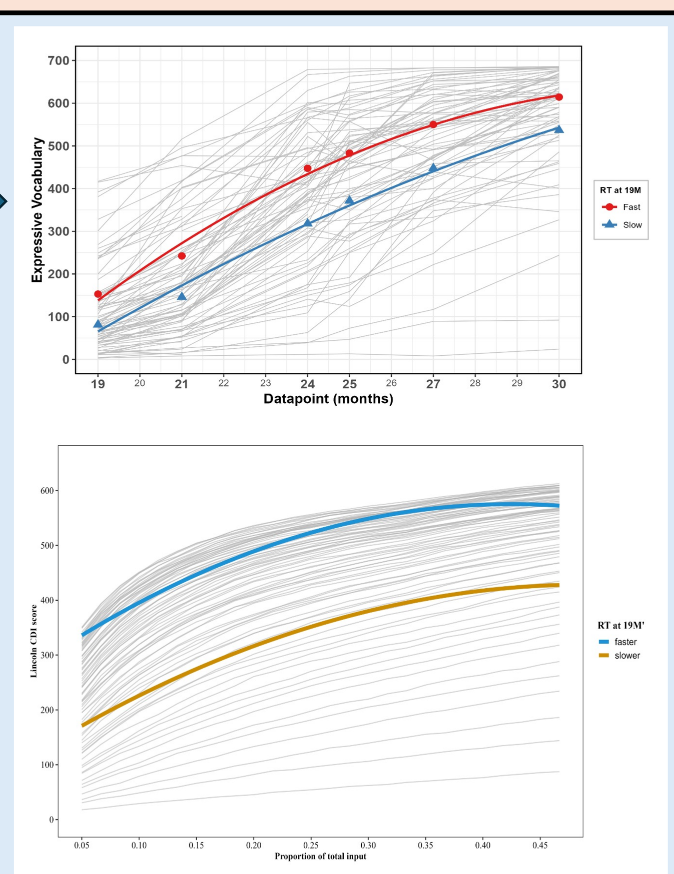
Larger vocabulary = more likely to represent familiar words with a smaller number of chunks: **/cat/** instead of **/c/ /a/ /t/**

Each chunk takes the same time to process regardless of length, so the fewer chunks needed, the faster the processing speed.

A larger store of sub-lexical chunks = more likely to learn new words more quickly. This is because new words can be encoded in fewer exposures.

Figure 1: Rowland et al. (in press)

Children



CLASSIC

Study 2: Eye-Tracking

Study 2 Aim: To explicitly and directly test the predictions of CLASSIC about the time it will take children to process different words.

Predictions:

1. Words that are represented by fewer chunks in CLASSIC (**few-chunk words**) will be processed more quickly than words represented by more chunks (**many-chunk words**).
2. Variance in processing speed will be smaller for few-chunk words than many-chunk words, because we expect that all/most children should be able to represent these words in fewer chunks.
3. The correlation between processing speed and vocabulary size will be smaller for few-chunk words versus many-chunk words, because these words are likely to be represented in fewer chunks for all/most children.

Study 2: Methods

Participants: 61 23- to 25-month-olds (M age = 23.35, SD age = 0.48).

- All participants were monolingual British English children.

Design:

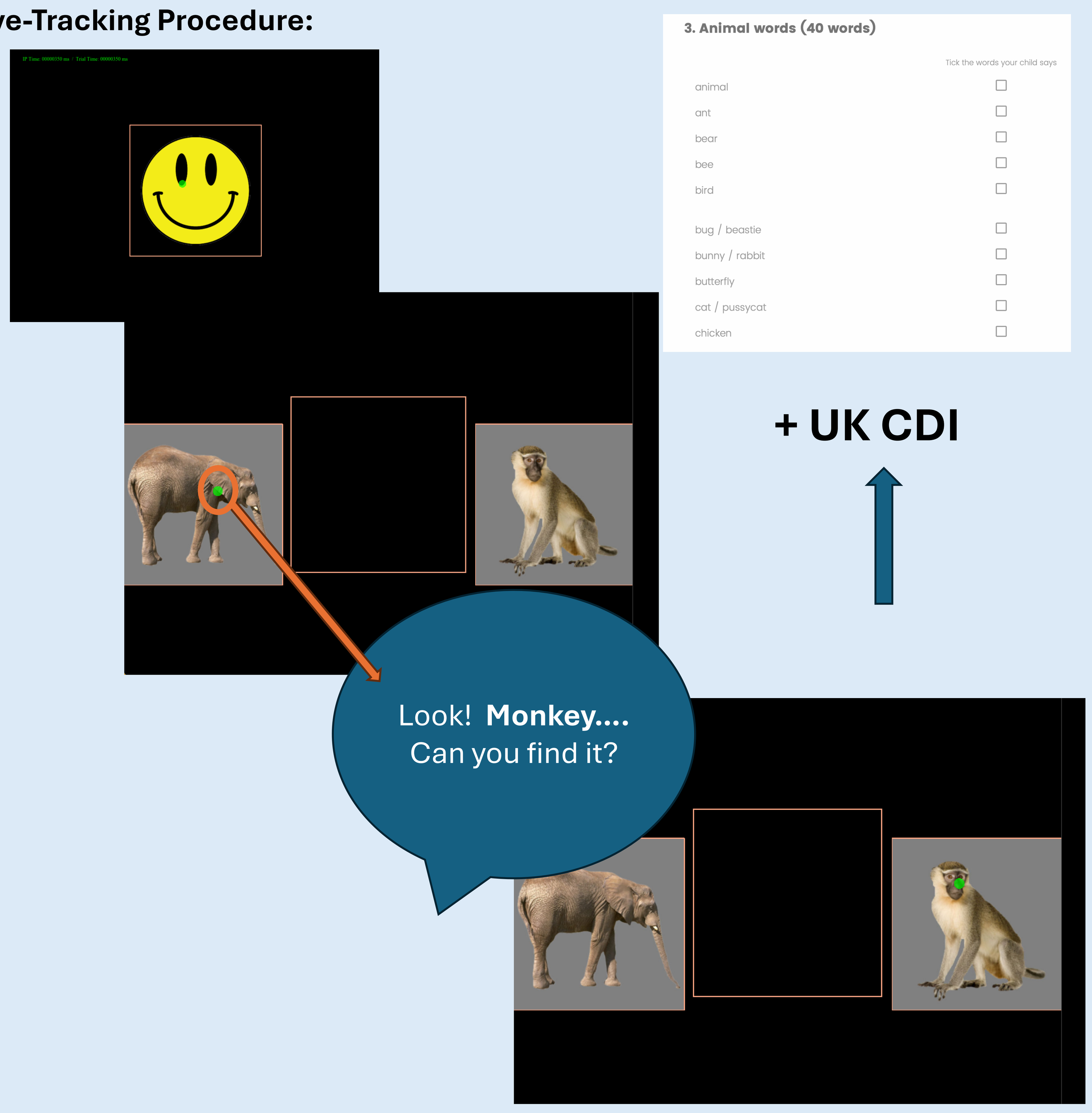
- Gaze-triggered looking while listening paradigm (Egger et al, 2020).
- 8 words CLASSIC needed only a few chunks to process (**<1.5 chunks**), and 8 which CLASSIC needed more chunks to process (**>1.8 chunks**).
  - All words were chosen from the Lincoln CDI and were balanced for length and familiarity (according to Wordbank norms for UK children).

Few-Chunk Words	Many-Chunk Words
Bee *1.11 chunks, 72% familiar, 1 syllable*	Bug *1.94 chunks, 73% familiar, 1 syllable*
Sock *1.48 chunks, 72% familiar, 1 syllable*	Fork *1.85 chunks, 71% familiar, 1 syllable*
Doll *1.32 chunks, 63% familiar, 1 syllable*	Frog *1.81 chunks, 71% familiar, 1 syllable*
Bottle *1.36 chunks, 73% familiar, 2 syllables*	Ice Cream *2.00 chunks, 72% familiar, 2 syllables*
Lion *1.25 chunks, 64% familiar, 2 syllables*	Crayon *2.11 chunks, 65% familiar, 2 syllables*
Monkey *1.36 chunks, 75% familiar, 2 syllables*	Turtle *2.41 chunks, 66% familiar, 2 syllables*
Paper *1.25 chunks, 67% familiar, 2 syllables*	Towel *1.85 chunks, 63% familiar, 2 syllables*
Elephant *1.28 chunks, 67% familiar, 3 syllables*	Butterfly *2.53 chunks, 64% familiar, 3 syllables*
M chunks = 1.30 *M familiarity = 69.13%* *M syllables = 1.75*	*M chunks = 2.06* *M familiarity = 68.13%* *M syllables = 1.75*

Outcome variables:

- Processing speed = reaction time in milliseconds (ms).
- Vocabulary = score out of ~600 on UK CDI.

Eye-Tracking Procedure:



Study 2: Results

Data:

- 64 test trials per participant - each target word was presented 4 times.
- Each participant was required to have **2 useable latencies** in each condition.
  - Trials were excluded for reasons such as audible disruption.
- Total N = **1496 trials** (few-chunk = 732, many-chunk = 764).

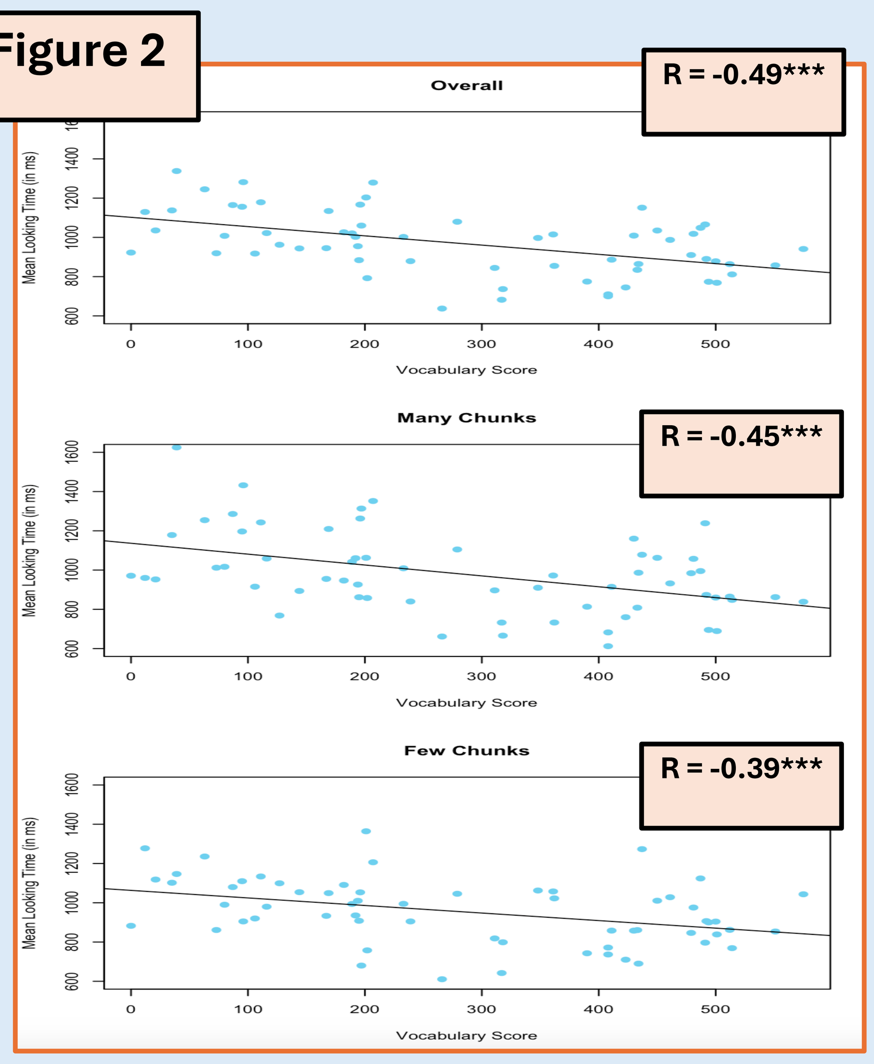
Results:

- **For prediction 1**: We show that there was no significant difference in processing speed between few-chunk words (M = 930.66, SD = 408.02) and many-chunk words (M = 946.87, SD = 437.42) (B = 12.07, SE = 46.04, p = .777).

  There was also no relationship when chunk was entered as a continuous measure (number of chunks) (B = 0.01, SE = 0.03, p = .738) nor when we removed words from the processing task that caregivers reported the children did not know (B = 13.08, SE = 46.32, p = .774).

- **For prediction 2**: The variance in processing speed was lower for few-chunk words ($\sigma 2$ = 166478.9) compared to many-chunk words ($\sigma 2$ = 191339.0), but this difference was not significant (F(763, 731) = 1.15, p = .058).
- **For prediction 3**: We show a significant negative correlation between vocabulary size and processing speed: higher vocabulary = faster processing. The size of the correlation was in the expected direction with a smaller correlation for few-chunk words (-.39) than many-chunk words (-.45) (see Figure 2).

Figure 2



Conclusion

- We find evidence for a significant relationship between vocabulary and processing speed (in line with Prediction 3).
- However, we find no evidence for a significant effect of chunk (not confirming prediction 1) nor a numerical difference in variance (not confirming prediction 2).
- Some possible explanations include:

1. Few-chunk and many-chunk words were balanced for **frequency** and **syllable length**. This means that the there was only a **small difference** in the number of chunks needed to process the words across conditions.
2. **Children** may have **more knowledge** than CLASSIC at this age meaning they could already be processing these words using very few chunks.

Future research could choose words 1) with **bigger differences** in the **number of chunks** and 2) that are **less frequent** = not already processed using only a few chunks.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57.
Egger, J., Rowland, C. F., & Bergmann, C. (2020). Improving the robustness of infant lexical processing speed measures. *Behavior Research Methods*, *52*, 2188-2201.
Gobet, F., Lane, P., Croker, S., Cheng, P., Jones, G., Oliver, I., & Pine, J. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236–243.
Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, *10*(6), 853–873.
Jones, G., Gobet, F., Freudenthal, D., Watson, S. E., & Pine, J. M. (2014). Why computational models are better than verbal theories: The case of nonword repetition. *Developmental Science*, *17*(2), 298–310.
Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352.
Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology*, *115*, 101238.
Rowland, C. F., Jones G., Peter, M. L., Bidgood, A., Jessop, A., Durrant, S., McLaughlin, P., & Pine, J. M. (in press). Explaining individual differences in children's vocabulary growth: Insights from the Language 0-5 project. *Language Learning*.

Scan here for poster.