Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

21 October 2025

Professorial Inaugural Lecture

The science and engineering of language acquisition in humans and machines

Prof Herman Kamper

Faculty of Engineering

#SUInauguralLectures

Hybrid Inaugural lecture of Prof Herman Kamper

Programme

Welcoming and introduction
Prof Wikus van Niekerk
Dean: Faculty of Engineering

Inaugural Lecture
Prof Herman Kamper
Electrical and Electronic Engineering

Q and A
Facilitated by Prof Herman Kamper

Rectorate Response
Prof Sibusiso Moyo
Deputy Vice-Chancellor: Research, Innovation and Postgraduate Studies

Vote of Thanks
Prof Japie Engelbrecht
Departmental Chair: Electrical and Electronic Engineering

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

21 October 2025

Professorial Inaugural Lecture

The science and engineering of language acquisition in humans and machines

Prof Herman Kamper

Faculty of Engineering

#SUInauguralLectures

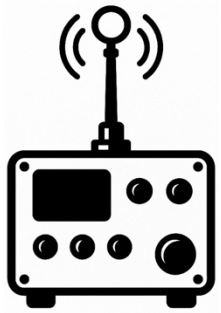
The Science and Engineering of Language Acquisition in Humans and Machines

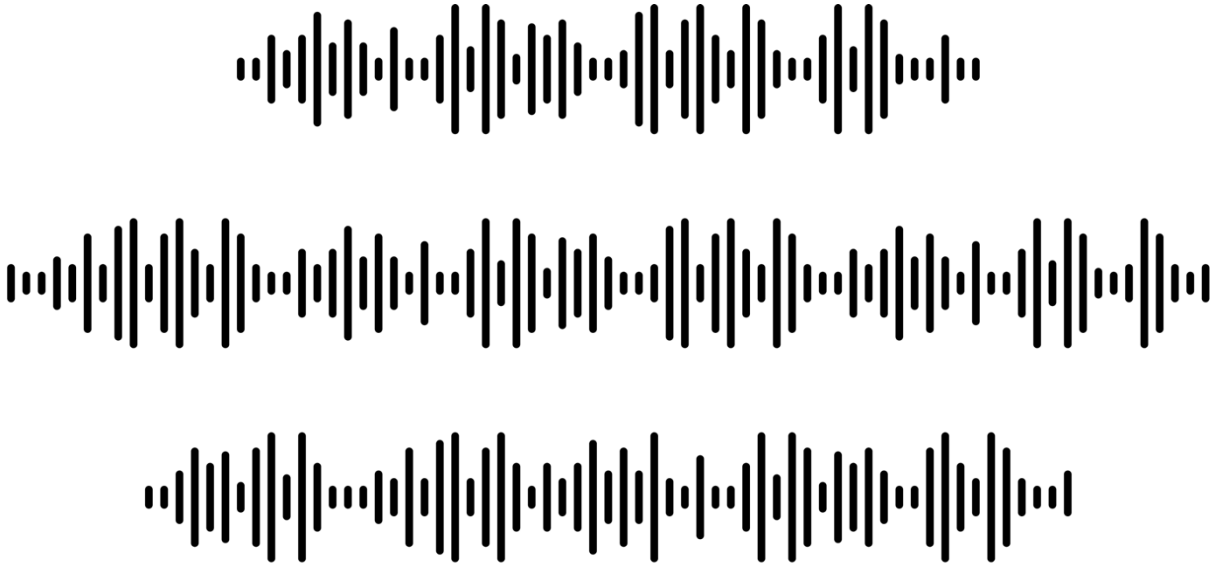
Herman Kamper

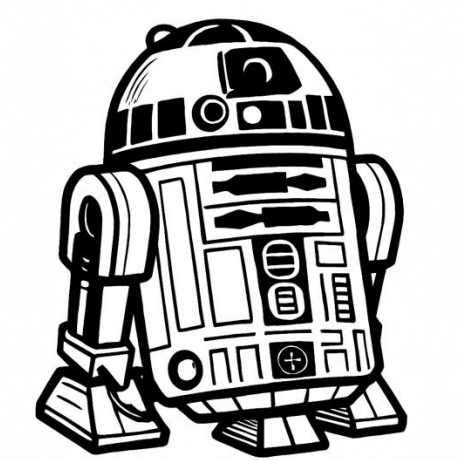
Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

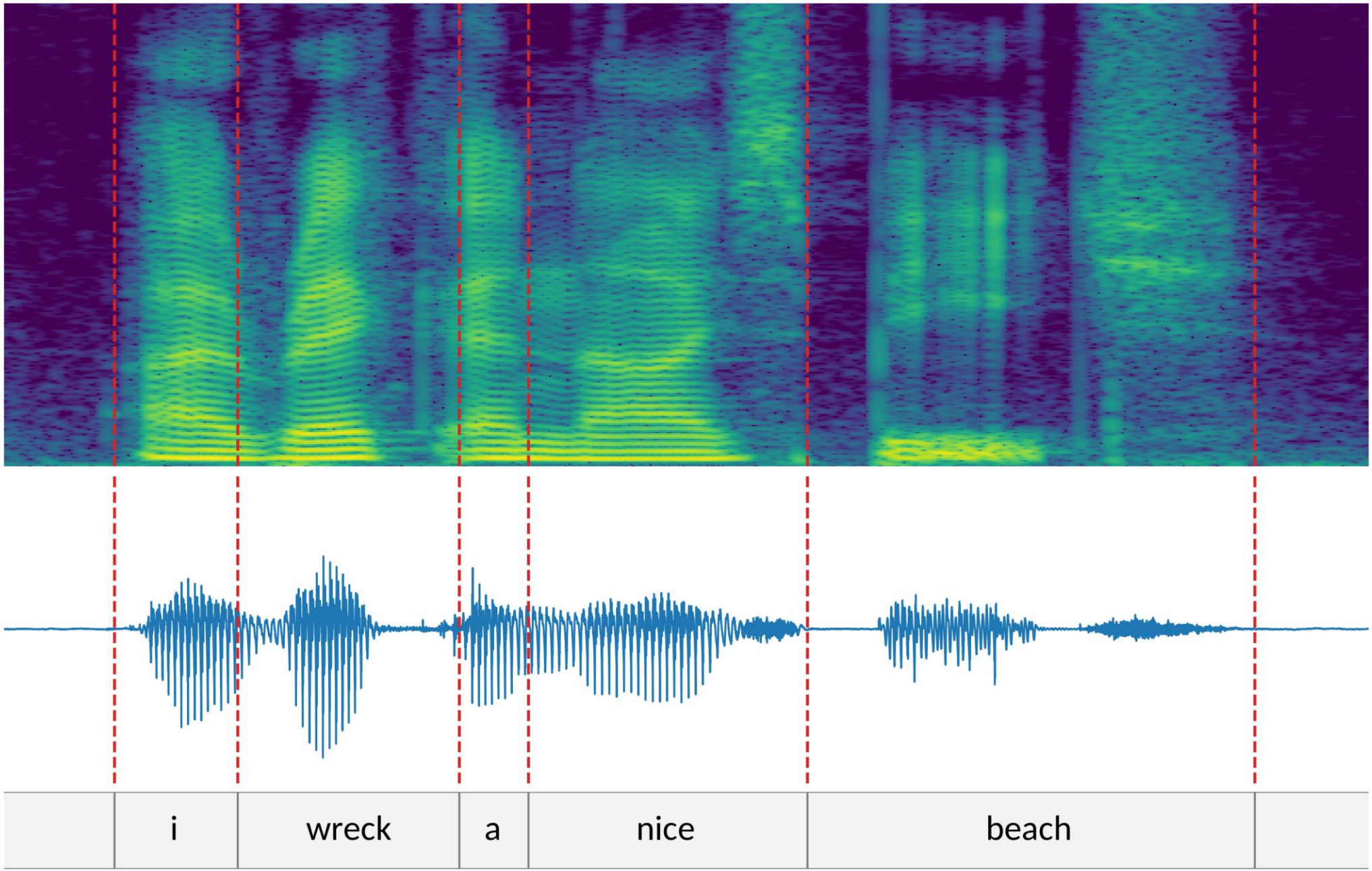
S. Singh, *The Code Book*, Fourth Estate, 1999.

Could you do this?





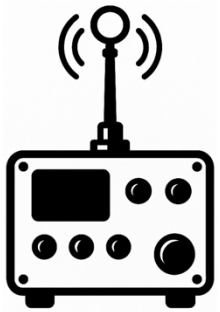


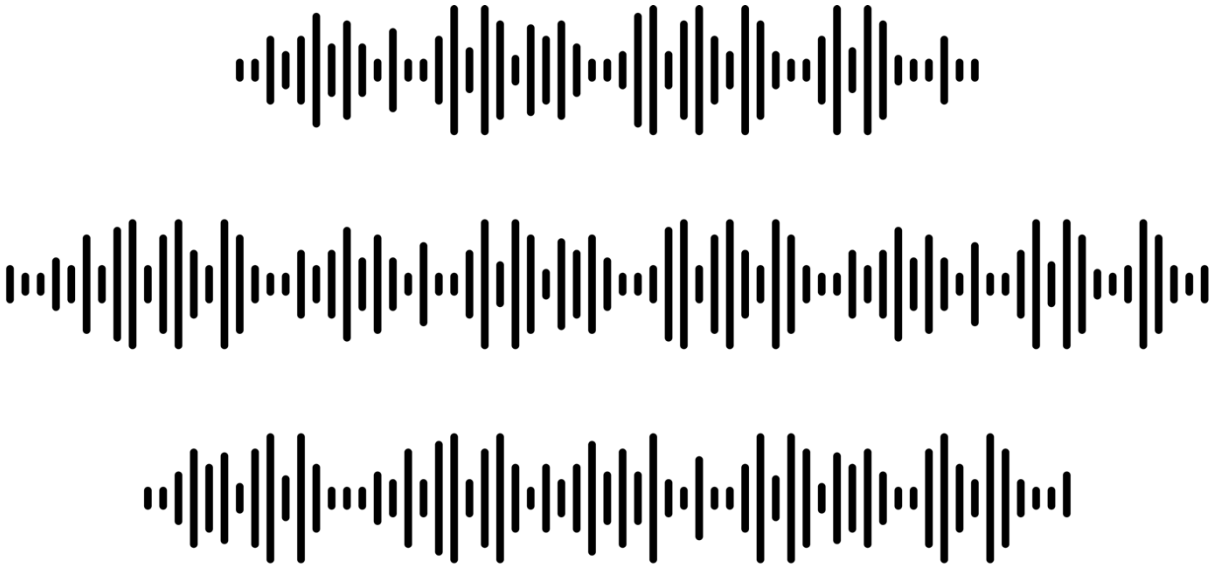

i
wreck
a
nice
beach

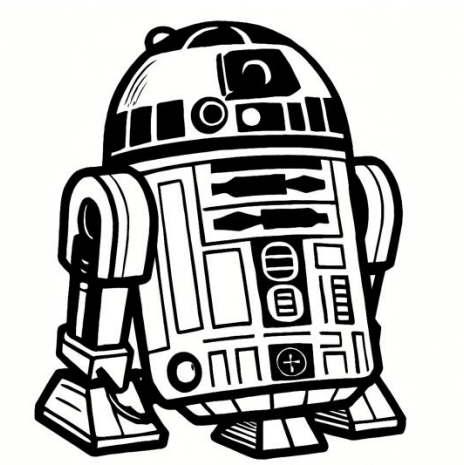
Can a machine do this?

a long time ago in a galaxy far far away

Unsupervised machine learning from speech







Maybe this is impossible?

The science and engineering of language acquisition in humans and machines

Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT

- **Science:** Understanding some observed phenomenon

- **Engineering:** Building something

Mimicking infant language acquisition

Why try to understand the science of language acquisition?

A 3-year-old human:

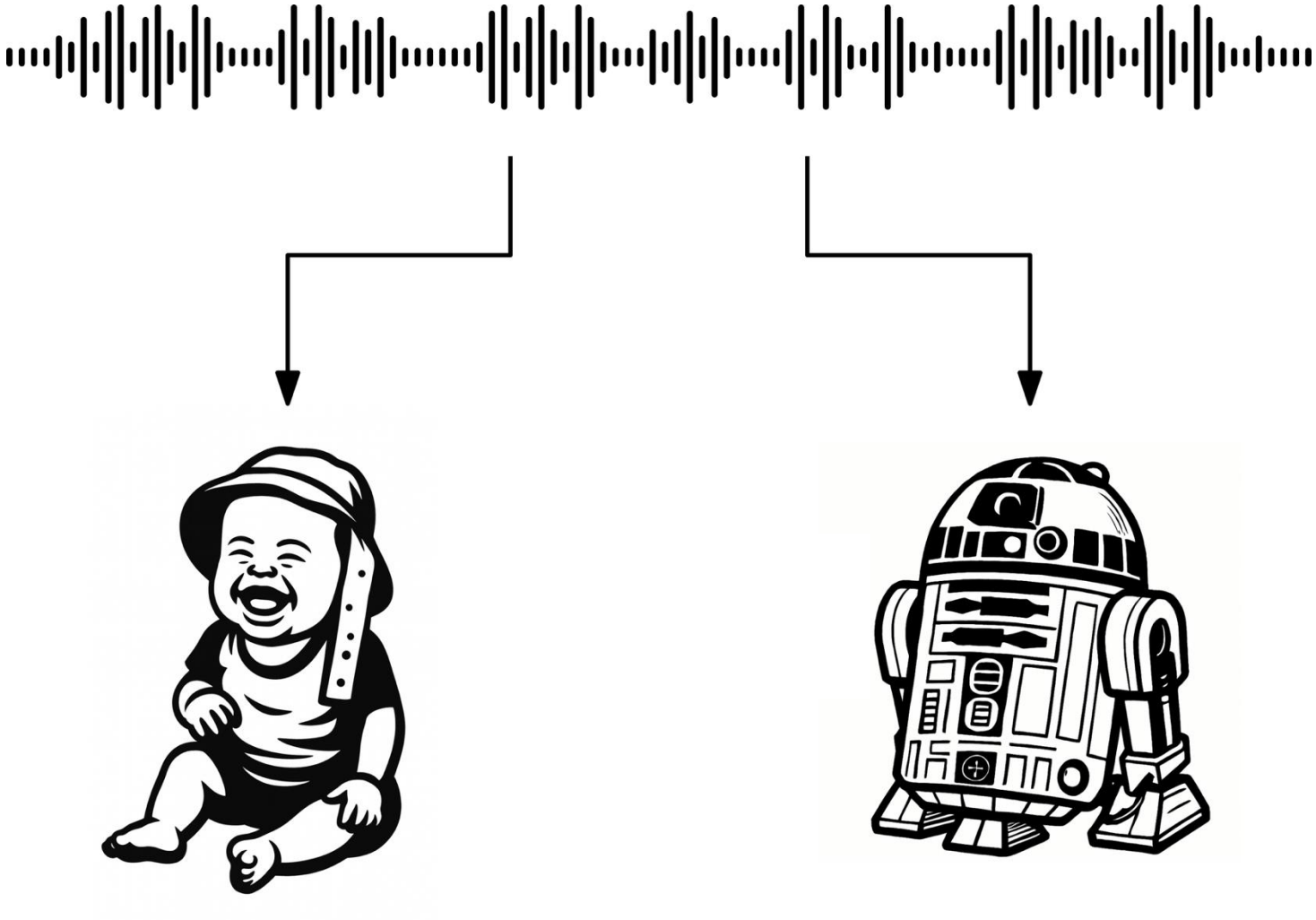
- 30 million words
- 3 years old
- Speech, no word boundaries
- Energy to train: 1.6 MWh

GPT-3:

- 100 billion words
- 10 000 years old
- Text, with word boundaries
- Energy to train: 1287 MWh

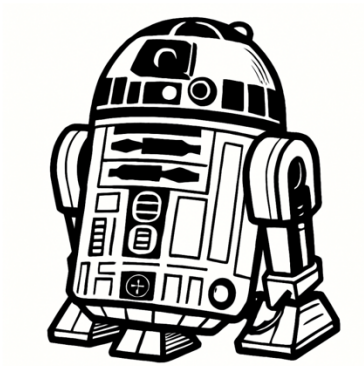
M. de Seyssel, “Introduction to language acquisition for speech processing researchers,” *Interspeech Tutorial*, 2025.
D. Patterson et al. “Carbon emissions and large neural network training,” *arXiv*, 2021.

Cognitive science: Reverse engineering



E. Dupoux, “Cognitive science in the era of artificial intelligence: A roadmap for reverse-engineering the infant language-learner,” *Cognition*, 2018.

	Infant study 1	Infant study 2	Infant study 3
Model 1			
Model 2			
Model 3			
Model 4			



COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13314

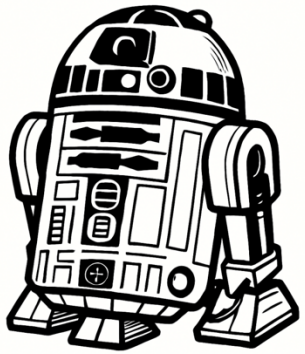
ISSN: 1551-6709 online
DOI: 10.1111/cogs.13314

Infant Phonetic Learning as Perceptual Space Learning: A Crosslinguistic Evaluation of Computational Models

Yevgen Matusevych,[a,b] Thomas Schatz,[c] Herman Kamper,[d] Naomi H. Feldman,[e,f] Sharon Goldwater[a]

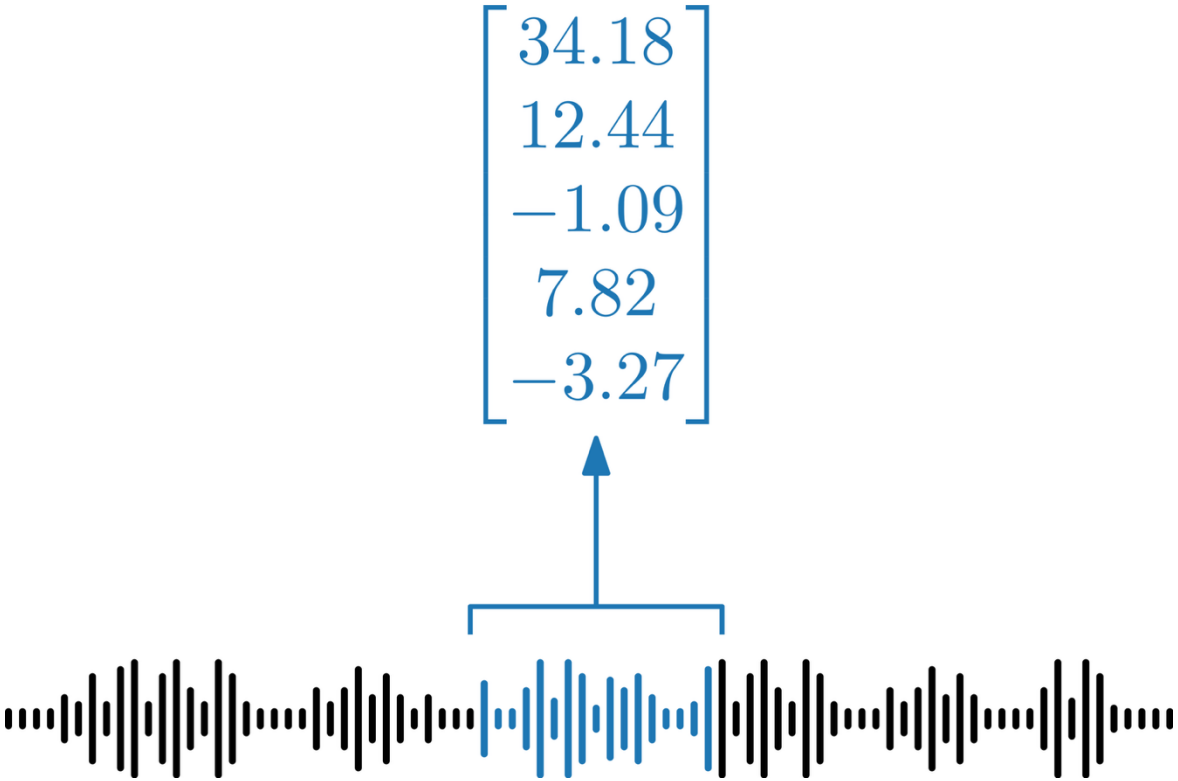
[a] *School of Informatics, University of Edinburgh*
[b] *School of Philosophy, Psychology and Language Sciences, University of Edinburgh*
[c] *CNRS, LIS, Aix-Marseille University*
[d] *Department of Electrical and Electronic Engineering, Stellenbosch University*
[e] *Department of Linguistics, University of Maryland*
[f] *Institute for Advanced Computer Studies, University of Maryland*

Received 29 September 2022; received in revised form 25 May 2023; accepted 19 June 2023

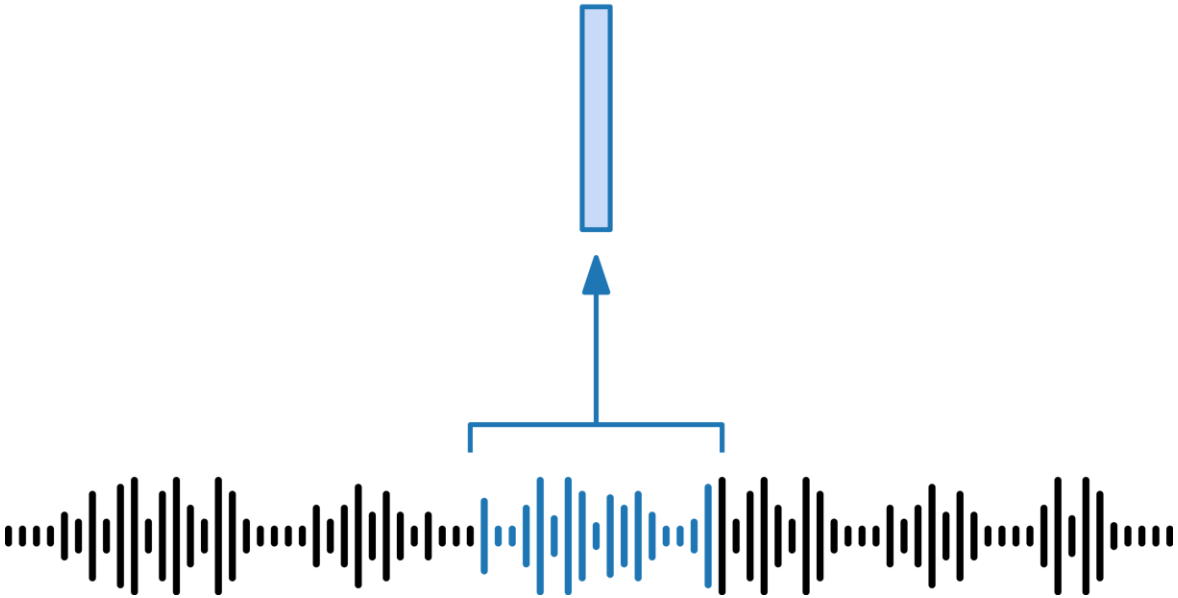


	Infant study 1	Infant study 2	Infant study 3
Model 1	✓	✓	✗
Model 2	✓	✗	✗
Model 3	✗	✗	✗
Model 4	✗	✗	✓

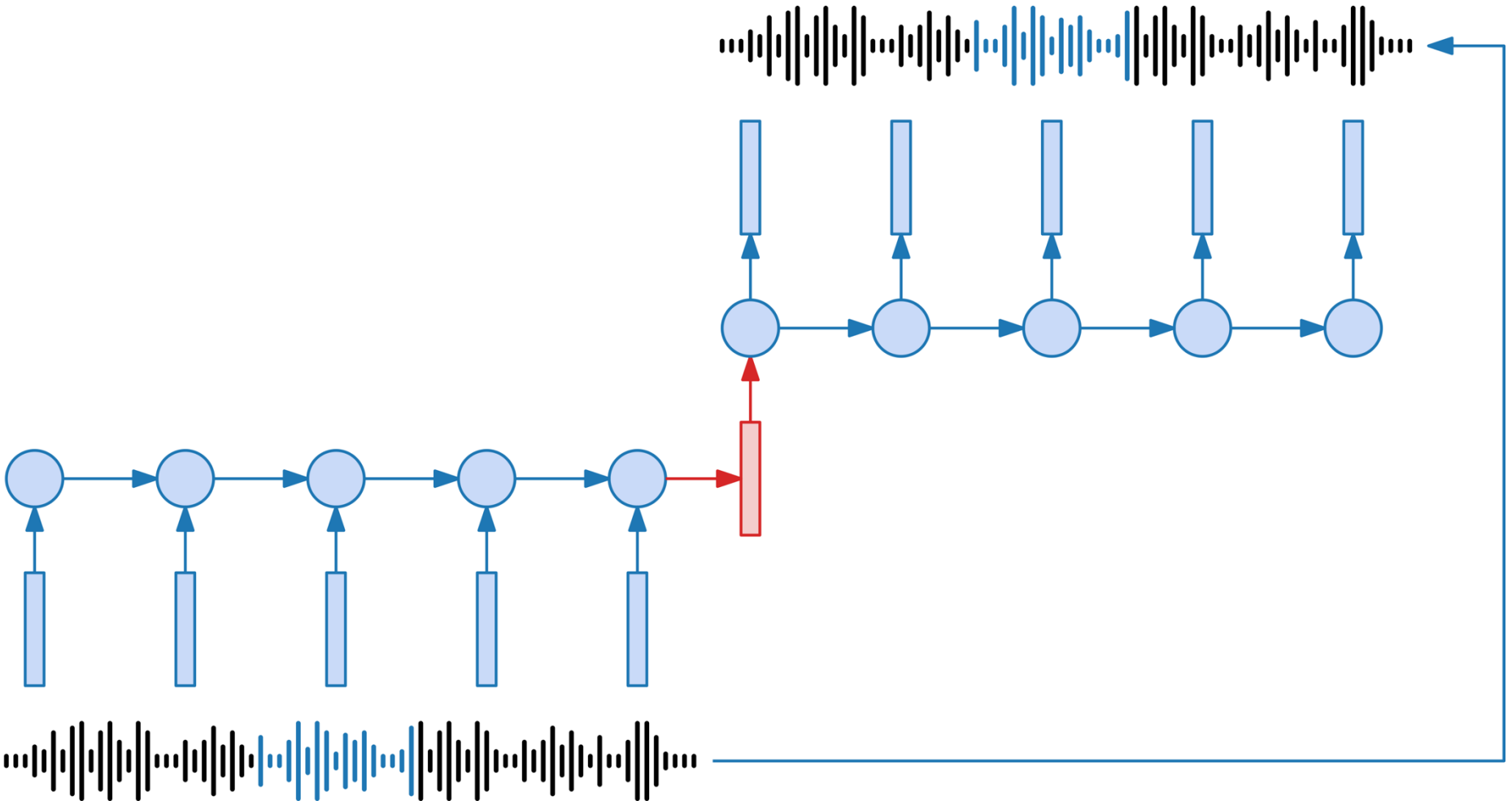
Models that learn without supervision

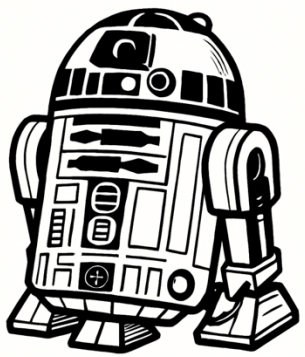


Models that learn without supervision



Models that learn without supervision





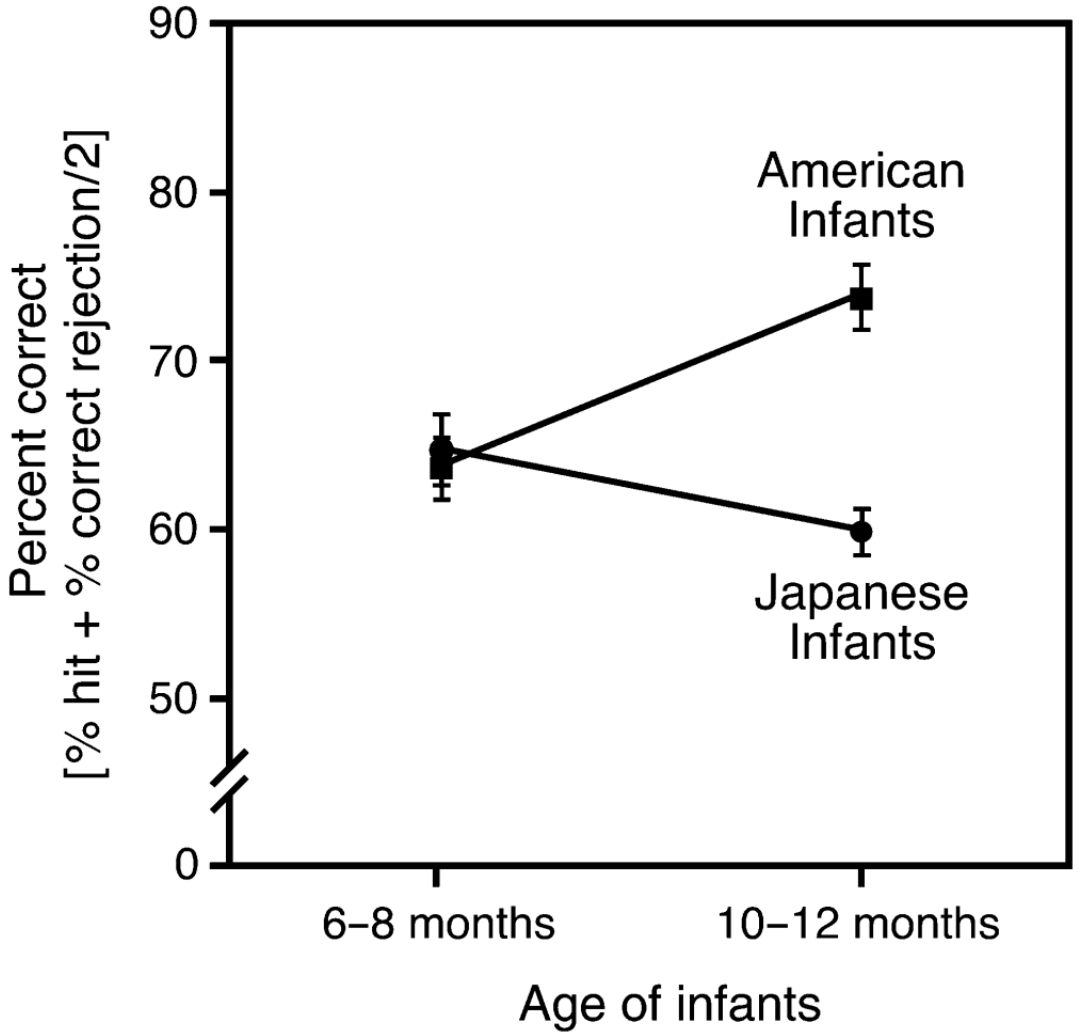
	Infant study 1	Infant study 2	Infant study 3
Model 1	✓	✓	✗
Model 2	✓	✗	✗
Model 3	✗	✗	✗
Model 4	✗	✗	✓

What do we know about infant phonetic learning?

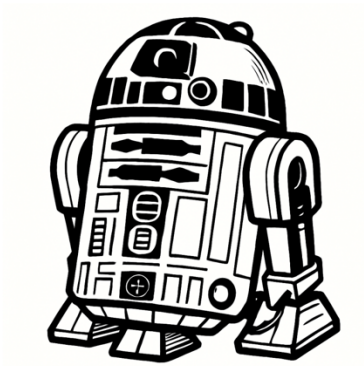
- English: [ɹ] ≠ [l]
  - rock – lock
  - wrong – long
- Japanese: [ɹ] ≈ [l]

- Mandarin: [ɕ] ≠ [tɕʰ]
- English: [ɕ] ≈ [tɕʰ]

- Catalan: [e] ≠ [ɛ]
- Spanish: [e] ≈ [ɛ]

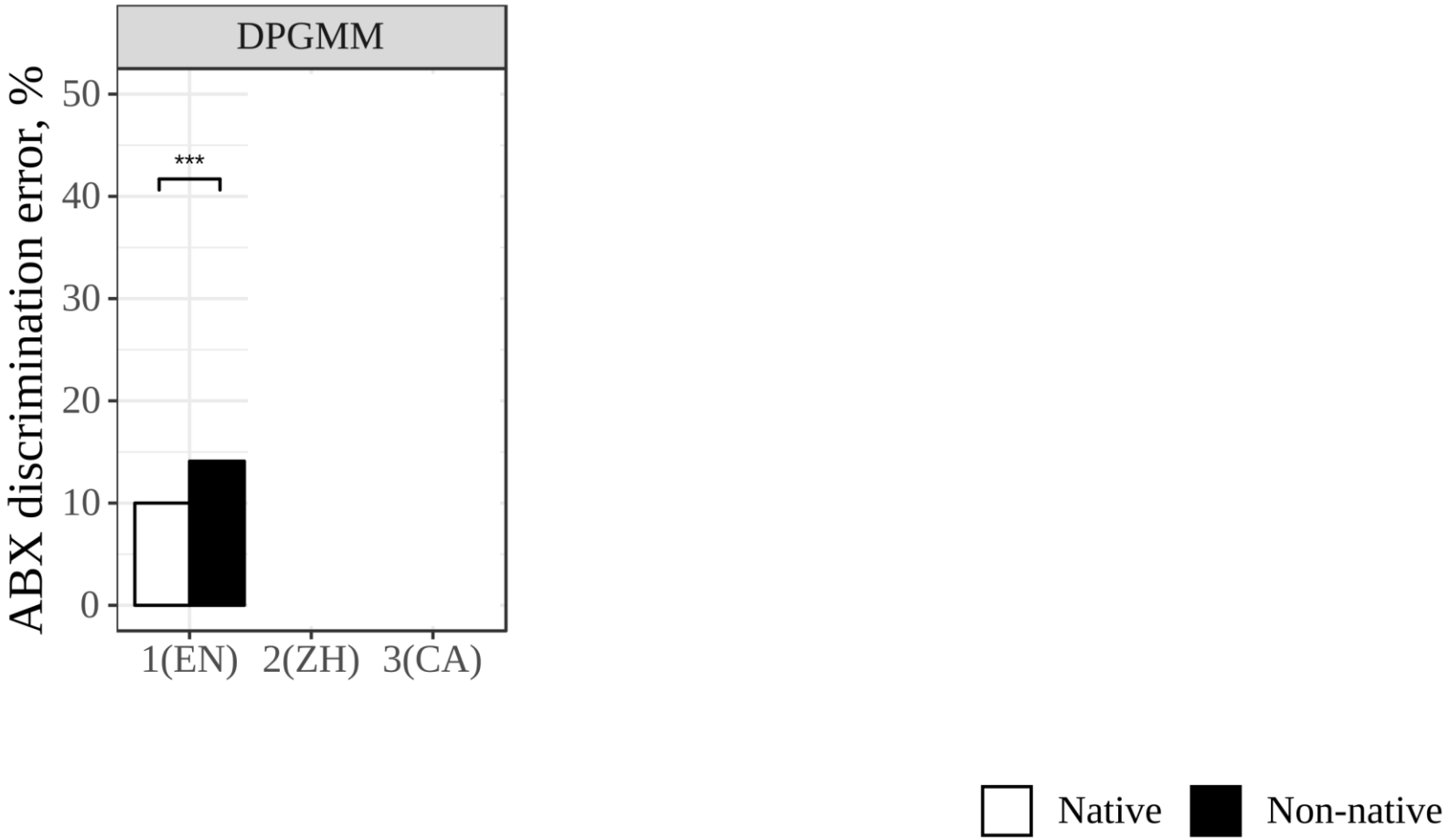


P. K. Kuhl et al., "Infants show a facilitation effect for native language phonetic perception between 6 and 12 months," *Developmental Science*, 2006.



	Infant study 1	Infant study 2	Infant study 3
Model 1	✓	✓	✗
Model 2	✓	✗	✗
Model 3	✗	✗	✗
Model 4	✗	✗	✓

Results: Models vs infants



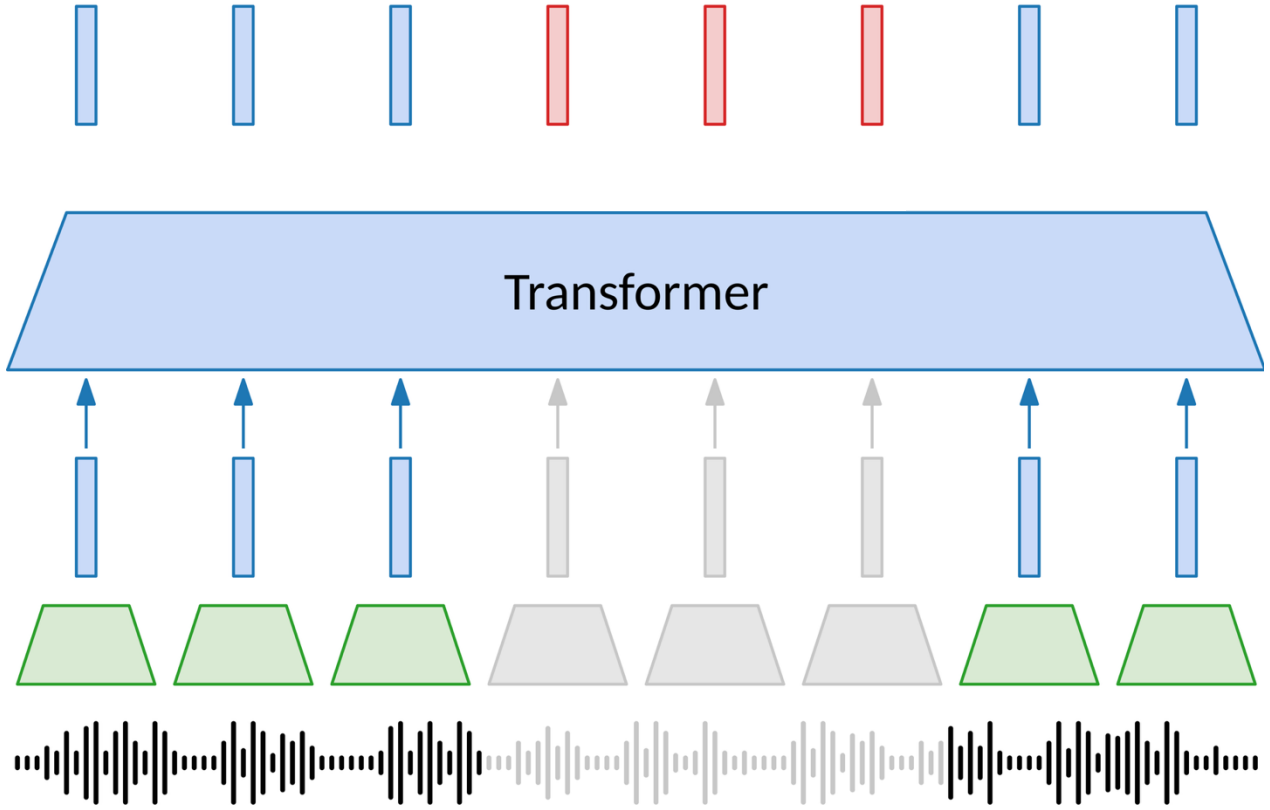
New hypotheses for infant testing

Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT

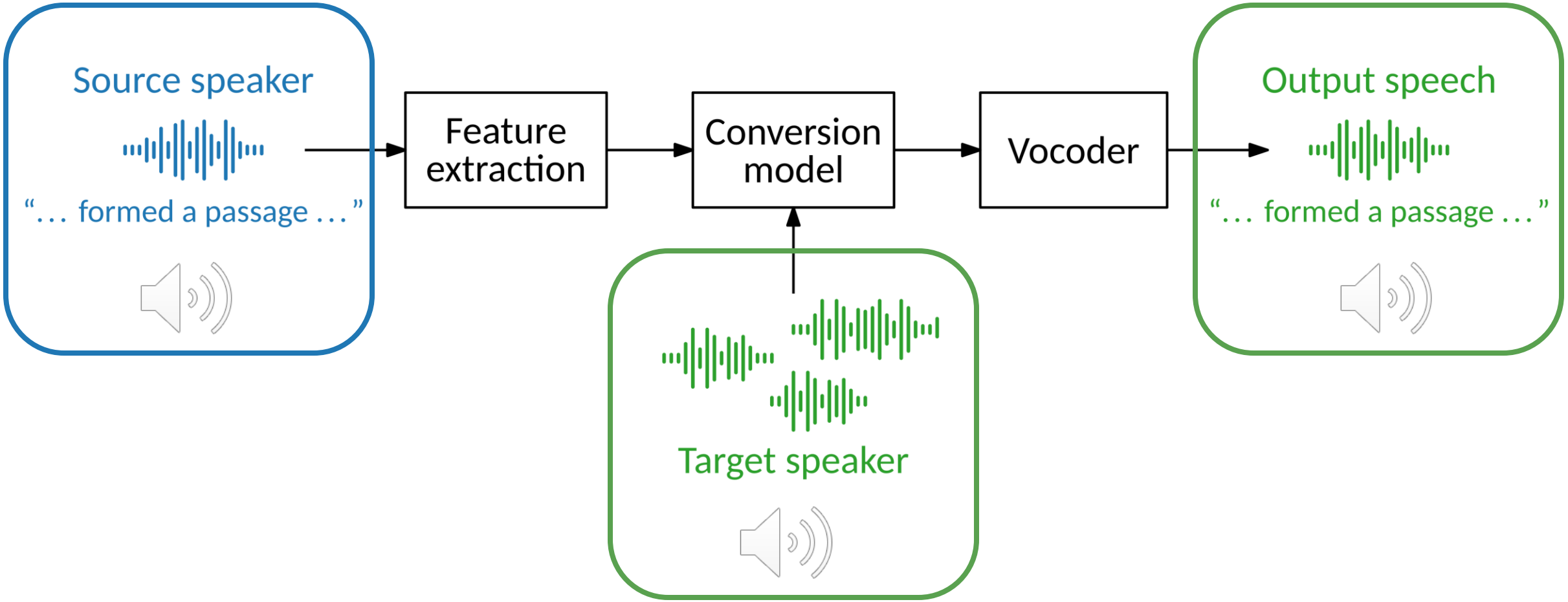
DPGMM		CAE-RNN	
Contrast	Mean difference	Contrast	Mean difference
[n]–[ɹ]	4.9	[f]–[z]	6.9
[d]–[ɹ]	4.9	[ʌ]–[ɔʊ]	5.8
[ɝ]–[l]	4.9	[f]–[s]	5.5
[ɝ]–[ɪ]	4.9	[l]–[ɹ]	4.8
[h]–[ɹ]	4.6	[m]–[ɹ]	4.5
[ʌ]–[ɝ]	4.4	[ɹ]–[w]	4.5
[m]–[ɹ]	4.5	[ʌ]–[ɑʊ]	4.3
[ð]–[ɹ]	3.8	[ɑ]–[ʌ]	3.0
[l]–[ɹ]	3.7		
[ɹ]–[v]	3.4		
[ɹ]–[t]	2.6		

Unsupervised models for new speech technology

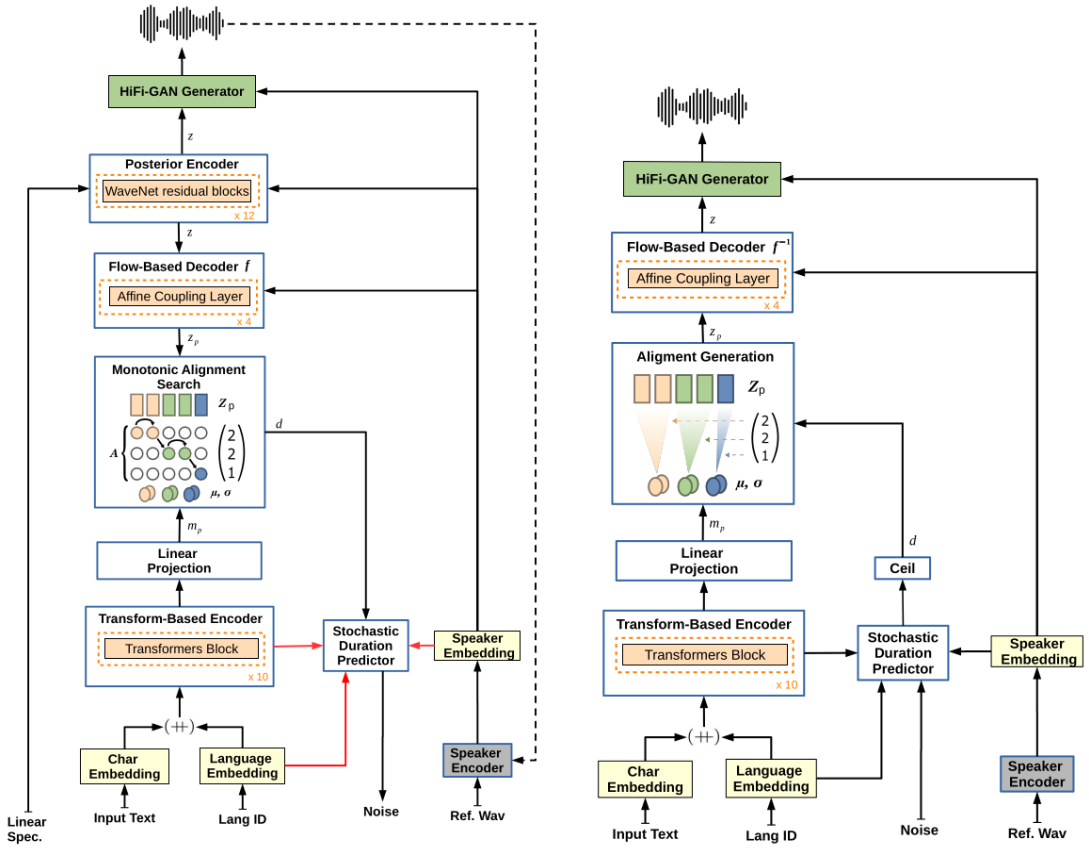
Predictive unsupervised speech model

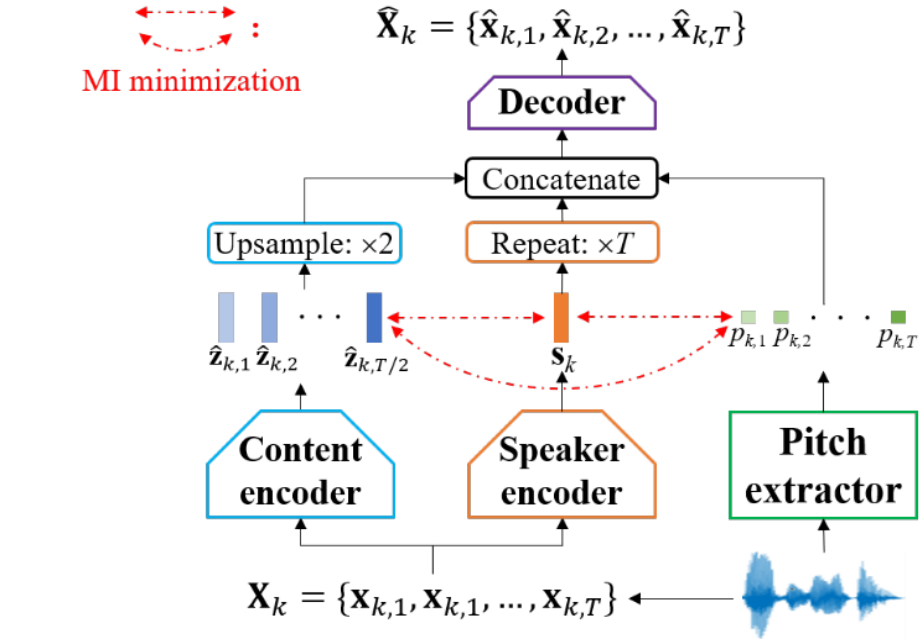


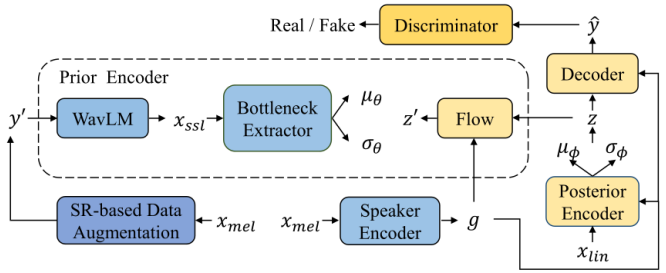
Voice conversion

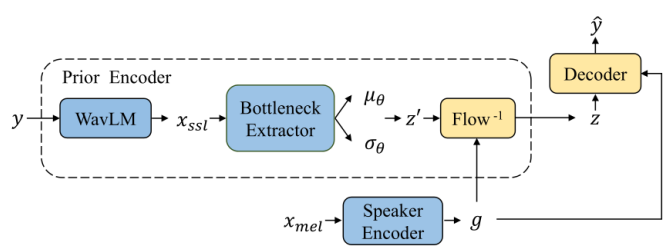


Previous voice conversion systems









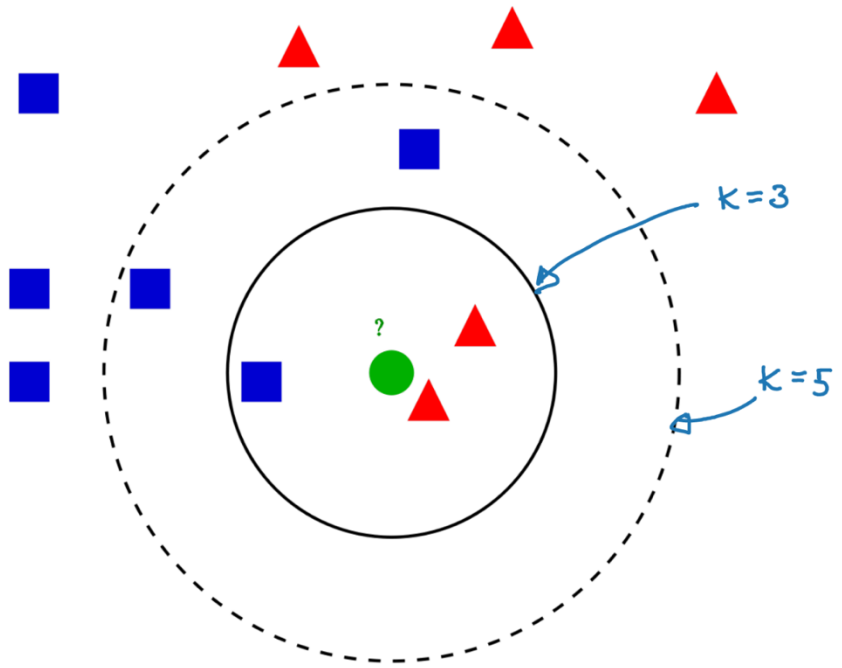
J. Li, W. Tu, and L. Xiao, "FreeVC: Towards high-quality text-free one-shot voice conversion," *IEEE ICASSP*, 2023.

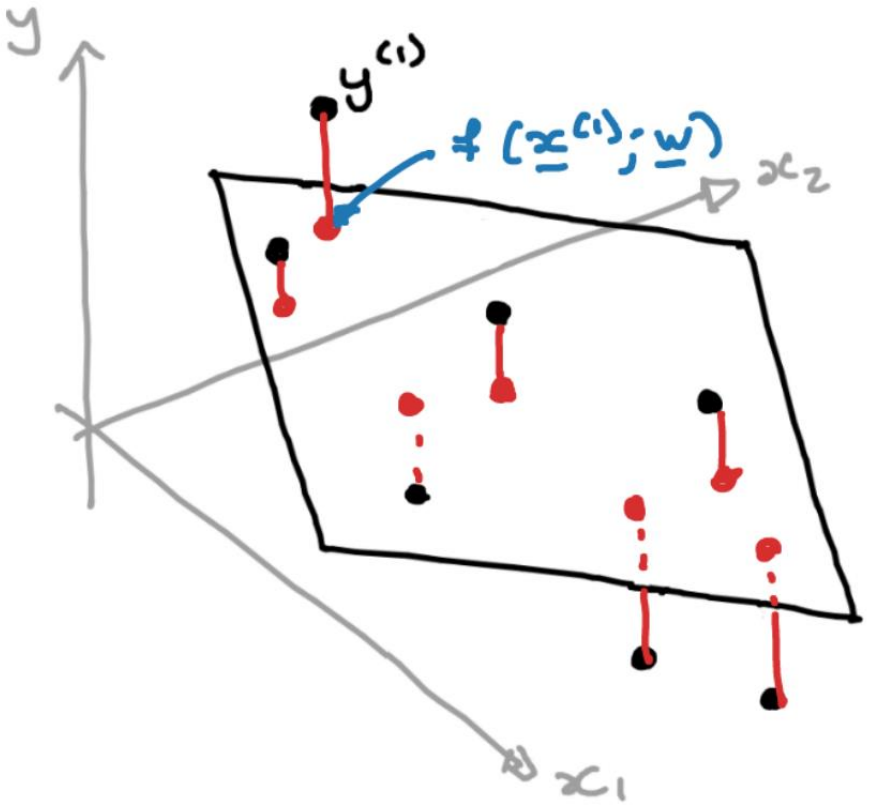
D. Wang et al., "VQMIVC: Vector quantization and mutual information-based unsupervised speech representation disentanglement for one-shot voice conversion," *Interspeech*, 2021.

E. Casanova et al. "YourTTS: Towards zero-shot multi-speaker TTS and zero-shot voice conversion for everyone," *ICML*, 2022.

Really simple machine learning on top of unsupervised speech representations

- K-nearest neighbours voice conversion (kNN-VC)
- Linear regression voice conversion (LinearVC)





Applications of kNN-VC

- Simple input and output example

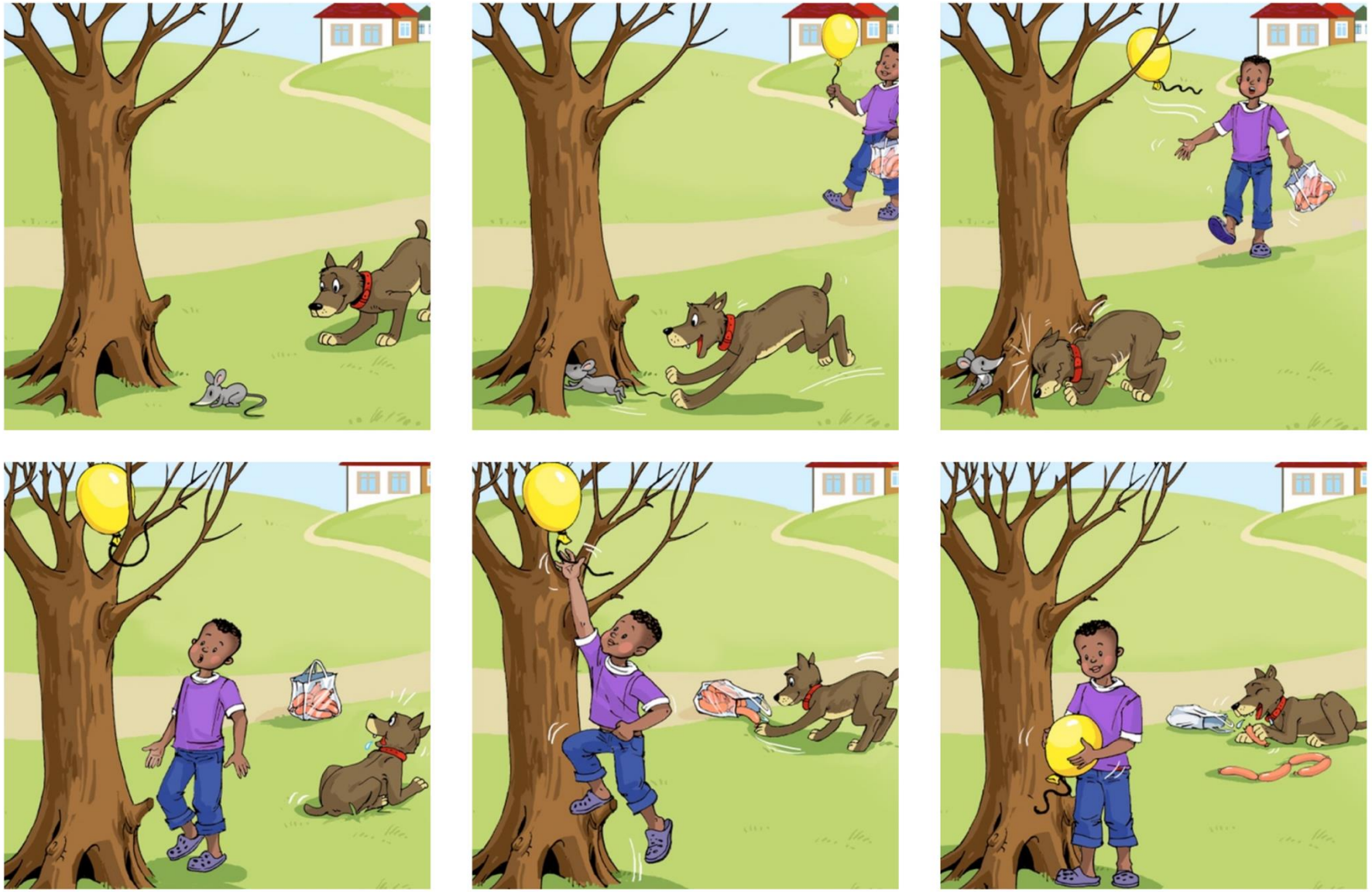
Output  Input 

- Cross-lingual for (bad) voice acting

Output  Input 

- Processing stuttered speech

Output  Reference 



A. Smith, “How narrative comprehension and production are intertwined with early learning indicators,” Master’s thesis, Stellenbosch University, 2023.
C. Jacobs et al., “Speech recognition for automatically assessing Afrikaans and isiXhosa preschool oral narratives,” *IEEE ICASSP*, 2025.

Summing up – food is close

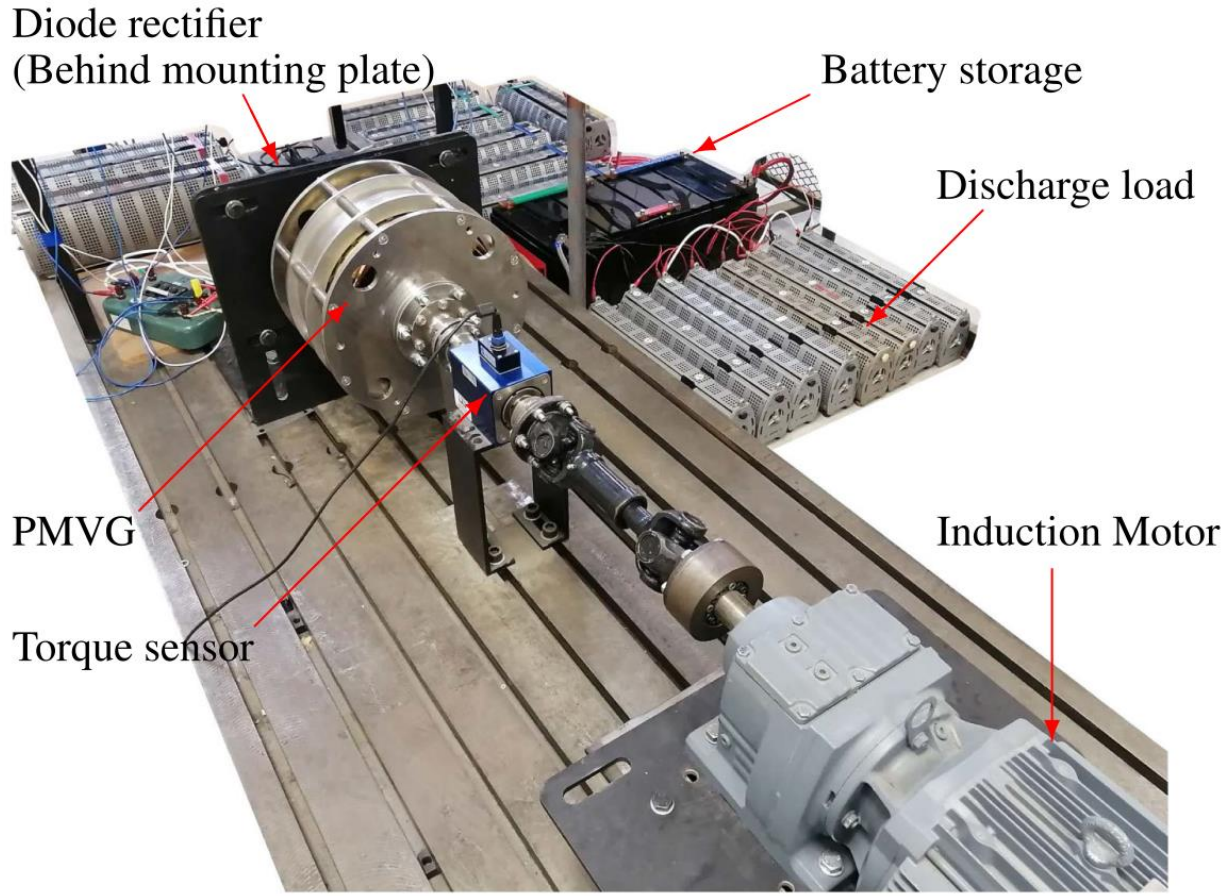
The science and engineering of language acquisition in humans and machines

Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT

- **Science:** Understanding some observed phenomenon (exploring creation)

- **Engineering:** Building something (shaping creation, stewarding, co-creating)


Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT
Thank you
Enkosi
Dankie


Diode rectifier
(Behind mounting plate)
Battery storage
Discharge load
PMVG
Induction Motor
Torque sensor


COMPAQ
COMPAQ

INTRODUCTION TO

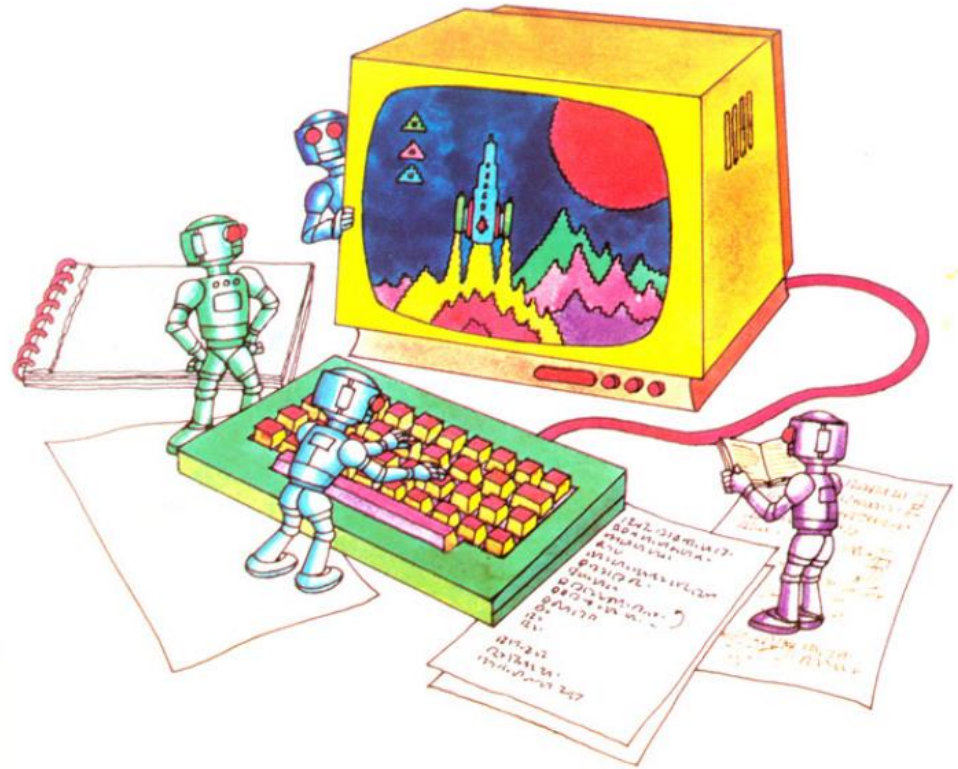
COMPUTER PROGRAMMING

Brian Reffin Smith

Edited by Lisa Watts

Designed by Kim Blundell

Illustrated by Graham Round
and Martin Newton



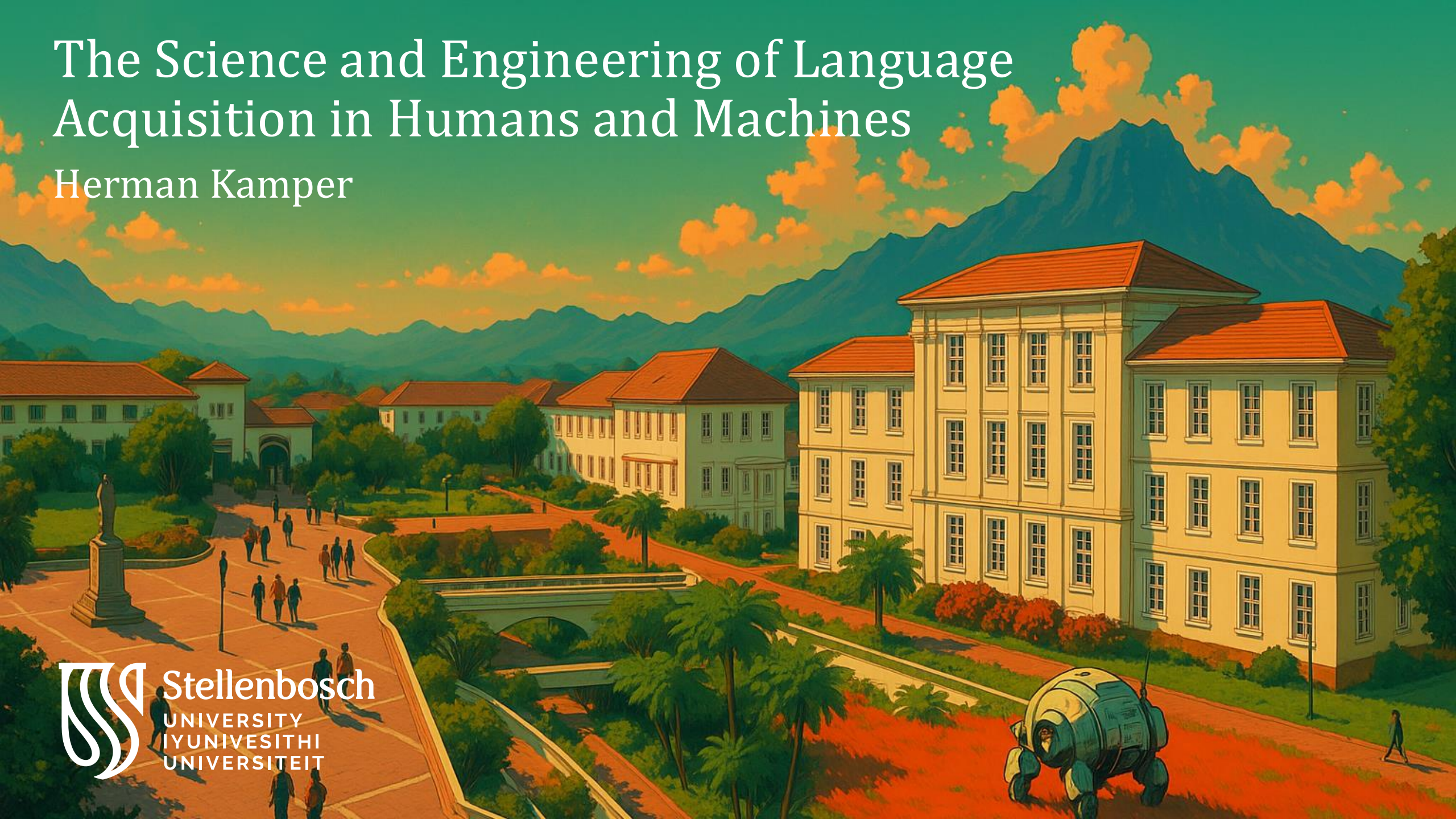
The Science and Engineering of Language Acquisition in Humans and Machines

Herman Kamper

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

The Science and Engineering of Language Acquisition in Humans and Machines

Herman Kamper



Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT