

# CEPOC: The Cambridge Exams Publishing Open Cloze dataset

Mariano Felice, Shiva Taslimipour, Øistein E. Andersen and Paula Buttery  
{mf501, st797, oa223, pjb48}@cam.ac.uk



## The open cloze test

### A bicycle you can fold up

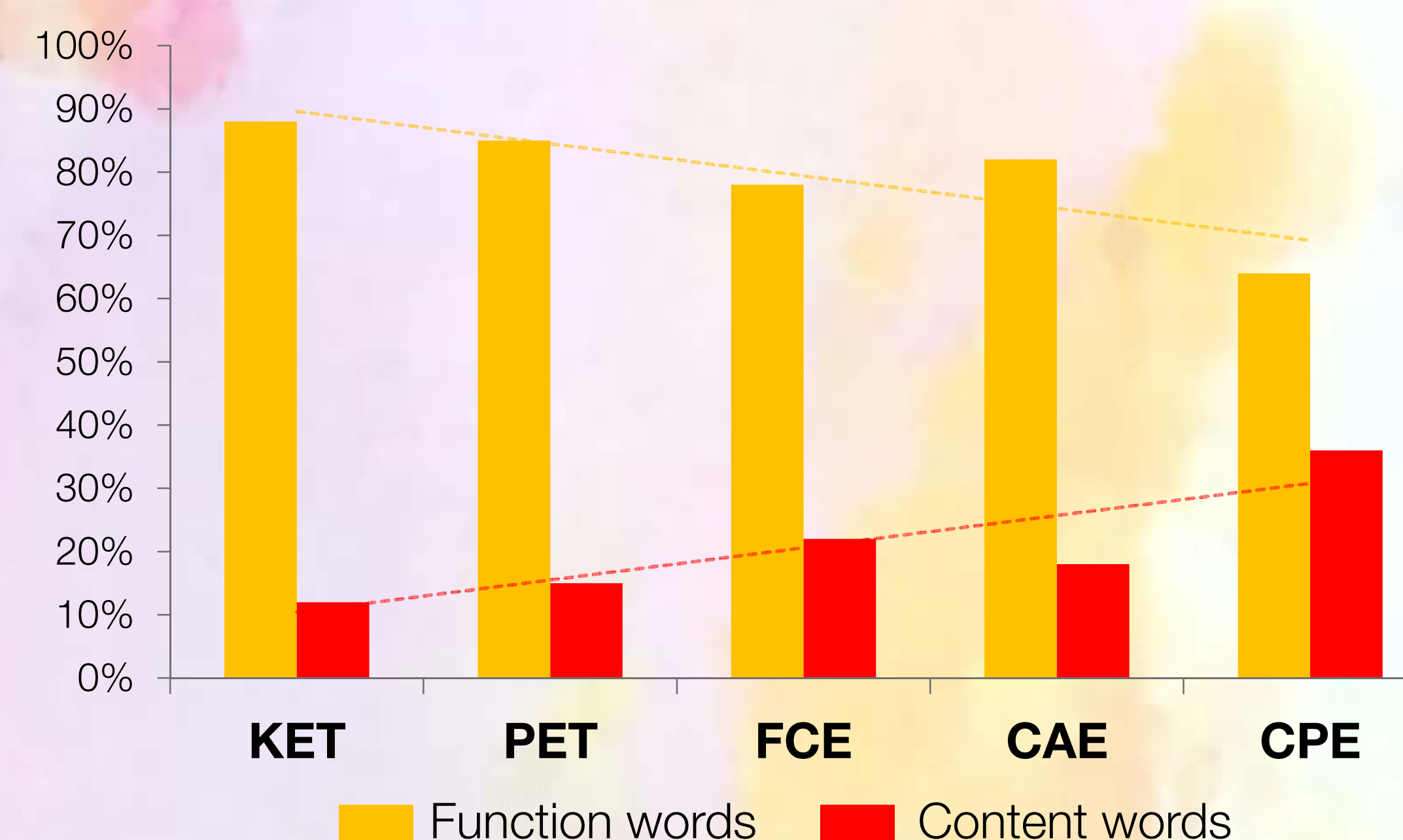
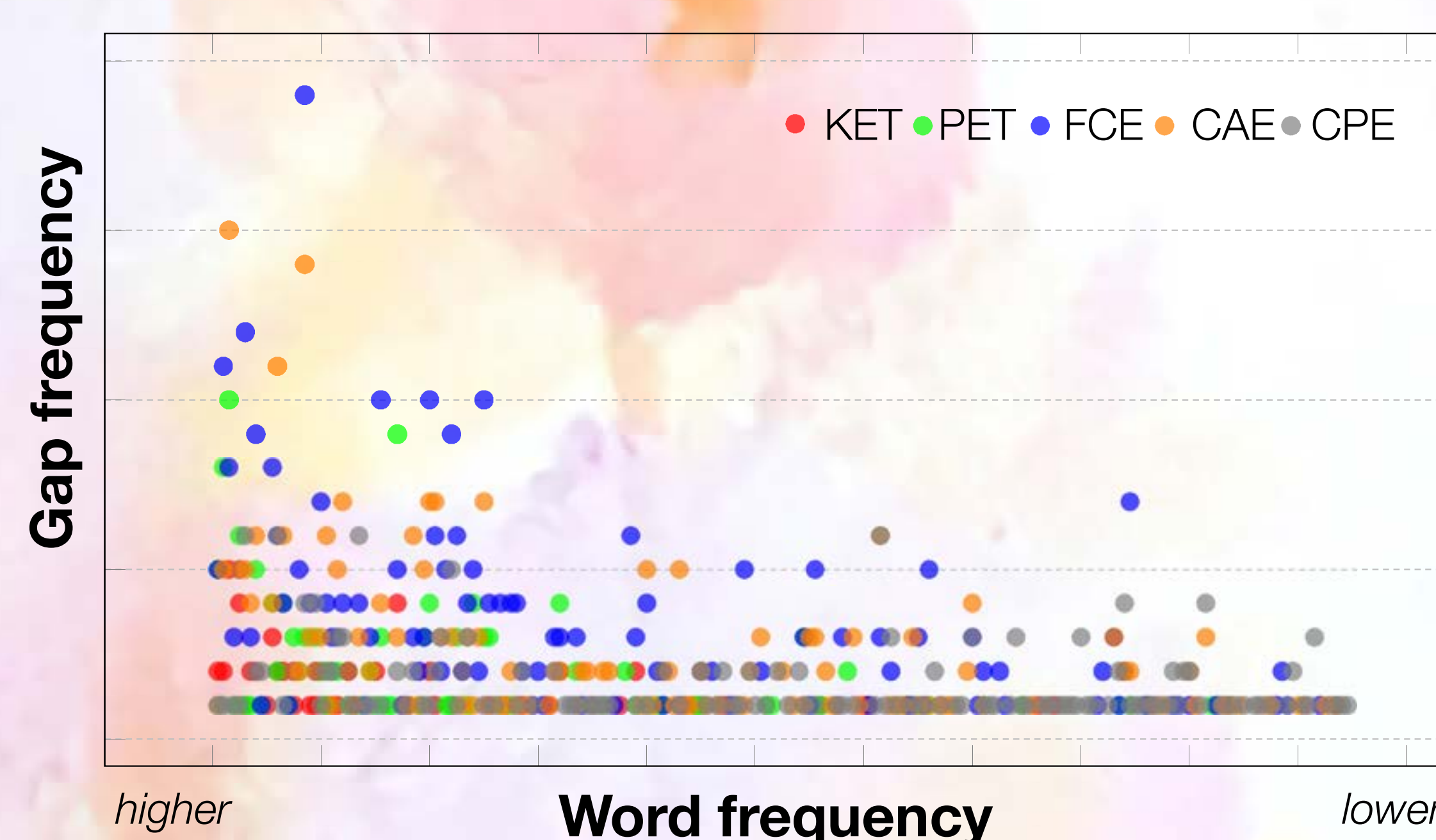
Folding bicycles have *been* around for quite some time now. However, an amazing new Japanese version \_\_\_\_\_ be folded with a swiftness and efficiency never seen before. This bike is designed \_\_\_\_\_ that it is possible to fold it up quickly. Once folded, you pull the bike along \_\_\_\_\_ ease.

## Dataset composition

Exam	CEFR level	#tests	Vocab. size	Avg. # tokens	Avg. TTR
Key (KET)	A2	6	290	128	65.21%
Preliminary (PET)	B1	21	854	163	61.91%
First (FCE)	B2	36	1,759	176	61.84%
Advanced (CAE)	C1	30	1,562	191	62.23%
Proficiency (CPE)	C2	21	1,270	192	63.56%

## Gap analysis

	KET	PET	FCE	CAE	CPE
# gaps per passage	11-12	6-8	9-11	9-13	9-11
# answers per gap	1	2	3	4	5
Total gaps	68	150	360	288	200
Avg. gap distance	10.81	19.36	17.74	19.23	19.49
Avg. gaps/sentence	0.97	0.69	1.15	1.07	1.14
Gaps at the beginning	27	57	116	111	66
Gaps in the middle	26	48	119	85	68
Gaps at the end	15	45	125	92	66



## EXPERIMENTS

### Gap filling

Model	KET	PET	FCE	CAE	CPE
bert-base-uncased	95.45	97.99	95.51	89.58	85.20
bert-large-uncased	94.03	98.00	97.77	92.01	90.36
roberta-base	98.53	<b>99.33</b>	96.38	94.79	93.50
roberta-large	<b>100.00</b>	<b>99.33</b>	<b>98.33</b>	<b>97.92</b>	<b>96.50</b>

Accuracy

### CEFR classification

<b>RoBERTa</b> Same training data as before. Texts filled in with first answer. Label with CEFR linked to exam.	<b>94.55</b>
<b>Random classifier</b> Predict 1 out of 5 classes.	20.00
<b>Majority class classifier</b> Predict all as B2.	31.57

Accuracy

### Gap prediction

**ELECTRA**  
Trained on 267 KET, 180 PET, 356 FCE, 281 CAE and 146 CPE open cloze tests.

**Random baseline**  
Based on gapped POS frequency in the training data.

**Exercise Maker**  
Rule-based system based on the most frequently gapped words in CUP&A's exams (Malafeev, 2014).

Model	KET	PET	FCE	CAE	CPE
Random baseline	23.28	12.58	14.95	14.79	11.00
Exercise Maker	-	-	23.17	15.28	10.30
<b>ELECTRA</b>	<b>60.29</b>	<b>40.67</b>	<b>53.89</b>	<b>47.22</b>	<b>50.00</b>

Accuracy

## Download

CEPOC is freely available at:  
<https://github.com/CambridgeALTA/cepoc>