

APPENDIX

A. List of Words

The following table lists the 50 top ranking words for the Understood dataset and the Produced dataset. Hebrew transliteration and English translation are given.

Ranking	Understood - DS_u		Produced - DS_p	
	Hebrew transliteration	English translation	Hebrew transliteration	English translation
1	Mzlg	fork	llcht	go
2	svvn	soap	shmchh	blanket
3	lshvvr	to break	mchnsm	pants
4	vrz	water tap	chtzr	yard
5	mchnsm	pants	rgl/rglm	leg/legs
6	gzr	carrot	chvs/sfl	cup/mug
7	vrz	rice	chys	chair
8	chnvt	shop	lchm	bread
9	chvlh	sick	chlvn	window
10	chvchv/chvchvm	star/stars	shmch	happy
11	chlvn	window	lshmvr	preserve/take care
12	chys	a chair	fshr/mvtr	allowed
13	msfrm	scissors	tzmch	plant
14	shfn/rnvt	rabbit	chvltzh	shirt
15	shvlchn	table	mdrgvt	Stairs
16	mtvch	kitchen	l'zvr	help
17	chvltzh	shirt	chvr/chvrh	friend
18	ktn	little	yd/ydm	hand/hands
19	lshmvr	preserve	l	on
20	frfr	butterfly	dlt	door
21	mlvch	salty	tytv/chytl	nappy/diaper
22	tmvnh	image	mvrst shynm	toothbrush
23	mgvt	towel	mtvch	kitchen
24	shvvr	broken	tzrch	must
25	ktzr	short	chlv	milk
26	vfnm	bicycle	mnvrh	lamp
27	mtzch	forehead	tchnh	tehina
28	lshchk	to play	tzyfvr	bird
29	gdvl	big	ln'vl	lock
30	lrvtz	run	dv/dvv	bear
31	lch/lchm	to life	glh	wagon
32	chvr/chvrh	friend	mtvs/vvrn	airplane
33	lnhvg	to drive	kfh	coffee
34	fnm	face	tzfrd	frog
35	tvv	good	lchvvt	turn off
36	dg	fish	mgvt	towel
37	vrch/vrchm	knee/Knees	yn/ynm	eye/eyes
38	chnh/mvsch	parking/garage	yshvvn/syr	potty
39	chvth/vytzh	omelet/egg	f	nose
40	mnvrh	lamp	dvvrh	bee
41	gshm	rain	slvn	living room
42	ltzt	to go out	nk	clean
43	llcht	go on a trip	chmvd	cute
44	mhr	faster	l'shvt	pretend
45	tmvl	yesterday	svs	horse
46	chvr	mouse	lrvt/lhrvt	to see/show
47	svfrmrkt/mchvlt	grocery store	lkvm	getting up
48	zvzv	fly	tmvl	yesterday
49	syr	pot	lshchk	to play
50	chrvm	beads	chshv	now

B. Comparative Models Implementation

To benchmark our model's performance, we implemented two baseline models based on existing open-source IRT frameworks. The first, IRT-CAT, is a Computerized Adaptive Testing model grounded in Item Response Theory, which adaptively selects the most informative items for each participant to estimate ability efficiently. This implementation is adapted from code written by Grzegorz Krajewski, based on earlier work by P. Król on Polish language acquisition data. The source code used is available at [git@github.com:DianaSK/CAT-IRT.git](https://github.com/DianaSK/CAT-IRT). The IRT-CAT computes the ability score of the subject. For fair comparison across methods we map all outputs to CDI scores. Rather than using a linear mapping, a sigmoid function was fit to map the estimated abilities of the dataset to the CDI raw scores (see Figure 16). Let θ denote the estimated ability; the estimated scores are computed as follows:

$$\text{EstimatedScore}_{\text{understood}}(\theta) = \frac{428.803}{1 + e^{-1.6218(\theta - 0.0114)}} + 1.9013 \quad (5)$$

$$\text{EstimatedScore}_{\text{produced}}(\theta) = \frac{424.713}{1 + e^{-1.8922(\theta - 0.3288)}} - 1.1762 \quad (6)$$

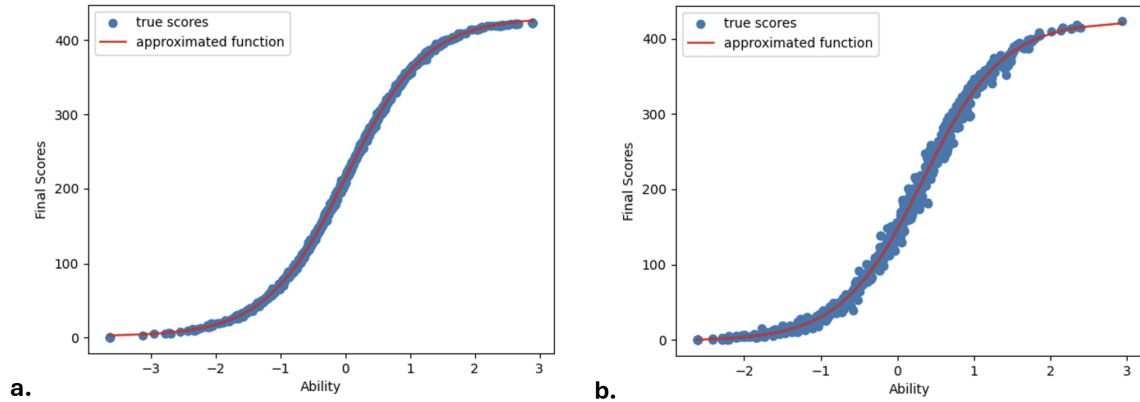


Fig. 16. Function fitting that maps estimated ability to CDI score of points (ranging between 0 and number of words in the full CDI-test).

The second model, IRT-CAT Bayes, employs a Bayesian approach to IRT parameter estimation using MCMC sampling. It is based on the implementation by the e-Babylab project (<https://github.com/lochhh/e-Babylab.git>) and modified for use in our study. Our adapted implementation is available at github.com/DianaSK/Baysian-IRT.