

Supplemental Material S1.

Table S1. Duration of auditory stimuli for Experiments 1 and 3.

Accurate word	Duration (ms)	Common substitute	Duration (ms)	Uncommon substitute	Duration (ms)
chick	560	ʃɪk	596	fɪk	587
leaf	653	wɪf	652	jɪf	610
thumb	575	fʌm	577	ʃʌm	569
comb	501	tʊm	507	pʊm	495
jar	520	dɑɪ	503	gɑɪ	451
safe	687	tɛɪf	633	pɛɪf	658
van	572	bæn	481	dæn	515
shirt	620	sɜ:t	561	fɜ:t	562
clock	555	kwɒk	541	kjɒk	548
rope	471	wɒp	519	jɒp	508
fish	623	pɪʃ	626	tɪʃ	631
girl	411	dɜ:l	395	bɜ:l	387
<i>M</i>	562		549		543
<i>SD</i>	79		74		77

Note. Table S1 shows stimuli used in Experiment 1 and Experiment 3. Accurate words are written orthographically, and misarticulated words are transcribed using International Phonetic Association transcription conventions. The average duration of each token, as recorded by an adult native speaker of American English, is listed to the right of each word in milliseconds. Mean and standard deviations for each condition are listed beneath each column.

Table S2. Nonobject picture characteristics.

Set 1				Set 2			
Kroll and Potter nonobject number	Child set size ^a	Strongest neighbor ^a	Object-likeness rating ^b	Kroll and Potter nonobject number	Child set size	Strongest neighbor	Object-likeness rating
nobj27	13	0.16	5.2	nobj88	15	0.07	3.2
nobj 75	9	0.09	3.3	nobj22	10	0.05	3.8
nobj61	11	0.08	5.7	nobj11	11	0.14	5.2
nobj58	10	0.16	3.1	nobj80	11	0.05	4.8
nobj29	11	0.20	3.3	nobj47	13	0.11	4.3
nobj01	15	0.05	3.6	nobj61	11	0.08	5.7
nobj53	14	0.08	3.4	nobj26	10	0.20	3.6
nobj81	10	0.05	5.1	nobj59	17	0.11	4.3
nobj31	9	0.11	5.2	nobj02	11	0.07	5.9
nobj13	16	0.05	3.6	nobj38	12	0.05	3.9
nobj17	11	0.03	5.7	nobj79	9	0.04	3.6
nobj52	12	0.11	3.4	nobj46	11	0.09	3.2
<i>M</i>	12	0.10	4.2	<i>M</i>	12	0.09	4.3
<i>SD</i>	2	0.05	1.1	<i>SD</i>	2	0.05	0.9
<i>Min</i>	9	0.03	3.1	<i>Min</i>	9	0.04	3.2
<i>Max</i>	16	0.20	5.7	<i>Max</i>	17	0.20	5.9

Note. Table S2 details the nonobject characteristics used as visual stimuli in Experiments 1–3. Nonobjects are from Kroll and Potter (1984). Each set is associated with one of two misarticulation types, and these are counterbalanced across versions of the study. “Strongest neighbor” is the highest neighborhood density of real words identified by children to associate with the nonobjects. Objectlikeness rating is a scalar rating of the objects “likeness” to real items. Mean, standard deviation, and range are listed beneath each set to demonstrate that each set was balanced. Each child saw all pictures twice, once with an accurate word, and once with a misarticulated word.

^aFrom Storkel, H. L., & Adlof, S. M. (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language, and Hearing Research*, 52(2), 306–320.

^bFrom Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, 23(1), 39–66.

Table S3. Counterbalancing of visual and auditory stimuli by block.

Block 1				Block 2			
Auditory stimulus	Condition	Response left	Response right	Auditory stimuli	Condition	Response left	Response right
chick	Accurate (Uncommon)	chick	nobj75	chick	Accurate (Common)	nobj79	chick
fick	Uncommon	nobj75	chick	shick	Common	chick	nobj79
thumb	Accurate (Uncommon)	thumb	nobj39	thumb	Accurate (Common)	nobj58	thumb
shumb	Uncommon	thumb	nobj39	fumb	Common	nobj58	thumb
jar	Accurate (Uncommon)	nobj29	jar	jar	Accurate (Common)	jar	nobj46
gar	Uncommon	jar	nobj29	dar	Common	nobj46	jar
van	Accurate (Uncommon)	nobj02	van	van	Accurate (Common)	van	nobj11
dan	Uncommon	nobj02	van	ban	Common	van	nobj11
clock	Accurate (Uncommon)	clock	nobj52	clock	Accurate (Common)	nobj47	clock
cjock	Uncommon	nobj52	clock	cwock	Common	clock	nobj47
fish	Accurate (Uncommon)	fish	nobj88	fish	Accurate (Common)	nobj01	fish
tish	Uncommon	fish	nobj88	pish	Common	nobj01	fish
leaf	Accurate (Common)	nobj81	leaf	leaf	Accurate (Uncommon)	leaf	nobj31
weaf	Common	leaf	nobj81	yeaf	Uncommon	nobj31	leaf
comb	Accurate (Common)	nobj26	comb	comb	Accurate (Uncommon)	comb	nobj22
tomb	Common	nobj26	comb	pomb	Uncommon	comb	nobj22
safe	Accurate (Common)	safe	nobj61	safe	Accurate (Uncommon)	nobj17	safe
tafe	Common	nobj61	safe	pafe	Uncommon	safe	nobj17
shirt	Accurate (Common)	shirt	nobj80	shirt	Accurate (Uncommon)	nobj38	shirt
sirt	Common	shirt	nobj80	firt	Uncommon	nobj38	shirt
rope	Accurate (Common)	nobj53	rope	rope	Accurate (Uncommon)	rope	nobj27
wope	Common	rope	nobj53	yope	Uncommon	nobj27	rope
girl	Accurate (Common)	nobj13	girl	girl	Accurate (Uncommon)	girl	nobj59
dirl	Common	nobj13	girl	bird	Uncommon	girl	nobj59

Note. Table S3 demonstrates the counterbalancing of blocks and nonobjects. This is one version of the study. Across other versions, the items that are “Response left” would be “Response right” and those in Block 1 would be in Block 2 and vice versa to ensure that no bias towards or against choosing any real objects or sets of real objects was apparent. Trials within a block were randomized by the experimental control software.

Table S4. Experiment 2 stimuli words, speakers, durations, and substitute.

Accurate	Speaker#	Duration (ms)	Substitute	Speaker#	Duration (ms)
chalk	pd 111	536	ʃak	pd 76	650
coat	pd 111	462	toot	pd 57	536
foot	pd 111	757	sot	pd117	686
giraffe	pd 89	817	dəɪæf	pd 98	678
goat	pd 99	332	doot	pd 117	367
kitchen	pd 89	921	tʃɪtʃən	pd 57	912
lady	pd 60	695	weɪdi	pd 89	728
ring	pd 89	440	wɪŋ	pd 76	542
shampoo	pd 91	1310	sæmpu	pd 98	786
soup	pd 91	839	ʃʊp	pd 111	1234
thief	pd 49	863	fɪf	pd 91	496
vacuum	pd 89	646	bækjum	pd 99	604
<i>M</i>		718	<i>M</i>		685
<i>SD</i>		265	<i>SD</i>		224
<i>Min</i>		332	<i>Min</i>		367
<i>Max</i>		1310	<i>Max</i>		1234

Note. Table S4 shows the stimuli used in Experiment 2. These were selected from previous recordings. Note that multiple child talkers with phonological disorders were used in both accurate and substitute conditions, and overall duration between the two conditions was controlled as closely as possible. Accurate words are written orthographically, and substitutes are transcribed using International Phonetic Association conventions.

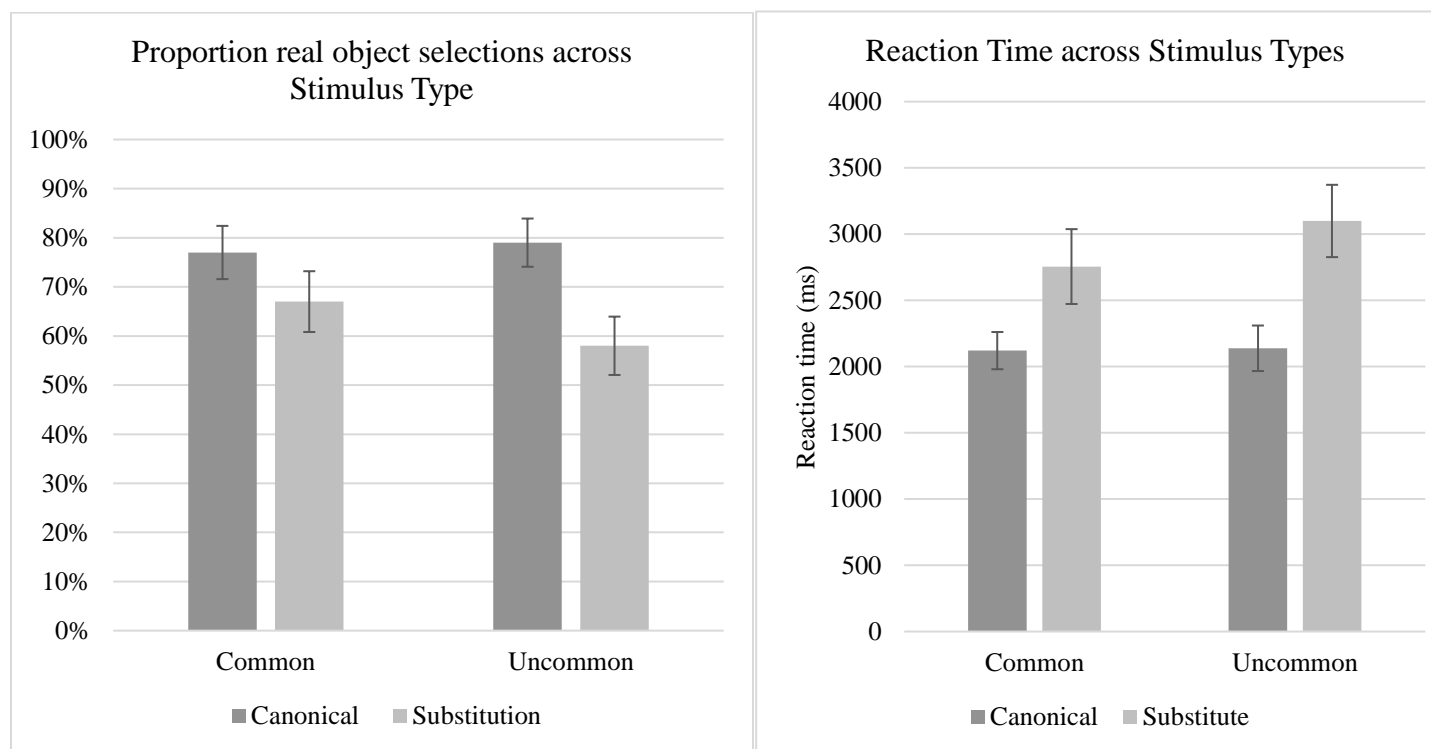
Experiment 3. Button box procedures and results.

Data Analysis

Accuracy and reaction time data were calculated in the same way as Experiment 1, but these data were calculated from the button box response method, rather than a mouse click as in Experiment 1.

Results

Figure S1. Mean proportion real object selections for button-box responses in Experiment 3(left). Mean reaction time for button-box responses across stimulus types for Experiment 3 (right).



The accuracy and reaction time results are displayed in Figure S1. In terms of accuracy (Figure S1, left), children chose real objects more for accurate productions than misarticulated productions, consistent with the results of Experiment 1 and Experiment 2. The proportion of real object selections were compared using a 2 Word Type (Accurate vs. Substitute) \times 2 Substitute Type (Common vs. Uncommon) repeated measures ANOVA. The main effect of Word Type was significant, $F(1, 14) = 9.77, p < .001, \eta_p^2 = 0.41$. No significant main effect of Substitute type was found, $F(1, 14) = 1.06, p = .32, \eta_p^2 = 0.07$. The interaction between Word Type and Substitute type was not statistically significant, $F(1, 14) = 4.05, p = .06, \eta_p^2 = 0.22$. Although the interaction was not statistically significant, we conducted post-hoc testing to result further examine similarities and differences between Experiment 1 and 3. We found no difference in real object selections between accurate productions ($M = 76\%$) and common

misarticulations ($M = 67\%$), $t(14) = 1.09$, $p = .09$, a finding that is inconsistent with results in Experiments 1 and 2. However, like Experiment 1 and Experiment 2, children chose real objects significantly more for accurate productions ($M = 79\%$) than uncommon misarticulations ($M = 58\%$), $t(14) = 3.54$, $p < .001$. Finally, real object selections for common misarticulations ($M = 67\%$) were not significantly greater than uncommon misarticulations ($M = 58\%$), $t(14) = 1.66$, $p = .12$. Recall that real object selections between common and uncommon misarticulation conditions were statistically significant in Experiment 1.

Mean reaction time (Figure S1, right) was compared using a 2 Word Type (Accurate vs. Substitute) \times 2 Substitute Type (Common vs. Uncommon) repeated measures ANOVA. A significant main effect of Word Type (Accurate vs. Substitution) was found, $F(1, 14) = 32.77$, $p < .001$, $\eta_p^2 = 0.70$. A significant main effect of Substitute Type (Common vs. Uncommon) was found, $F(1, 14) = 5.04$, $p = .04$, $\eta_p^2 = .27$. The interaction between Word Type and Substitute Type was not statistically significant, $F(1, 14) = 1.40$, $p = .26$, $\eta_p^2 = 0.09$. Planned post-hoc paired samples t tests were conducted to investigate the differences between accurate productions and misarticulations (Bonferroni correction, $p < .0125$). As in Experiment 1, children chose real objects significantly faster when hearing accurate productions ($M = 2120$) than common substitute productions ($M = 2754$), $t(14) = -3.39$, $p < .01$, and chose real objects significantly faster when hearing accurate productions ($M = 2137$) than uncommon substitute productions ($M = 3098$), $t(14) = -4.68$, $p < .001$. Subjects did not choose real objects significantly faster within substitution conditions: Common ($M = 2754$) vs. Uncommon ($M = 3098$), $t(14) = -1.76$, $p = .10$. Each of these findings is consistent with what was found in Experiment 1.

Discussion

Children selected real objects significantly more in accurate conditions than substitution conditions. They chose real objects more in accurate conditions than uncommon misarticulation conditions but not in the accurate vs. common misarticulation condition and not in the common vs. uncommon misarticulation comparison. These results are inconsistent with what was found in Experiments 1 and 2, where children consistently chose real objects significantly more for accurate productions than both misarticulation conditions. In terms of reaction time, much like the results of Experiment 1, children chose real objects significantly faster when hearing accurate productions than either common misarticulations or uncommon misarticulations. However, in accordance with previous findings, no reaction time difference between common and uncommon misarticulations was found. The lack of replication of Experiment 1 and 2 with Experiment 3 is likely a result of these analyses being underpowered due to low sample size. It is possible that with more subjects, and therefore more statistical power, the results may align more closely to the finding in our previous experiments. The results of this analysis trend in the same direction as previous findings in Experiments 1 and 2.

The difference in response method—mouse click vs. button box—produced a difference in the overall proportion of real object selections. Children selected real objects less frequently with the button box in Experiment 3 than with the mouse click in Experiments 1 and 2. For example, for correct articulations, children selected the real object on average 92%–94% of the time in Experiments 1 and 2 but only 77%–79% of the time in Experiment 3. This finding suggests that children made more errors in object selection with the button box (in Experiment 3) than with the mouse (in Experiments 1–2). Although the button mapping appears straightforward (i.e., left-picture = left button; right picture = right button), these results suggest that children

find mouse-clicking a more natural and intuitive method for selecting pictures than a button press. Therefore, where accuracy is of key relevance, a mouse-click response is an appropriate method for preschoolers, because there is a visual component that allows the child to check their response with the location of the cursor on the actual item to be selected.

However, if processing, such as response time, is of key importance, the length of time required to make a selection with a mouse may mask response time differences in preschool-aged children. Note that the reaction times in Experiment 3 with the button press were much shorter than in Experiments 1 and 2 with the mouse movement and mouse click. Likewise, as discussed in the main text of the article, eye-gaze results were the most sensitive measure for detecting processing differences between common and uncommon misarticulation conditions as opposed to mouse trajectory. Although the method for selecting objects may be clearer with mouse tracking, children's agility in moving the mouse may not be developed enough to support the use of mouse tracking as a precise method for measuring reaction time and response trajectory.