A communicative framework for early word learning

XXXXX and XXXXX

S1 Corpus Referential Talk and Gesture Coding Supplemental

Using the session transcripts, we coded all the concrete referents (in either speech or gesture) throughout the session that were produced by the parent or child. We further coded whether a referent was likely to be physically present, based primarily on transcript context. Coding was led by Madeline Meyers. A full coding manual can be found at **XYZ.COM**

ADD FINAL LINK FOR FULL CODING MANUAL.

S1.1 Coding A Referent

In our coding, we aimed to capture any referential communication about concrete in gesture or speech. A concrete noun can roughly be thought of as any noun that could be pointed to if it was present. Note that a signle utterance may have multiple referents. For gesture, any deictic gesture toward an identifiable referent definitionally refers to a concrete noun. For references in spoken language, note that the referent sometimes differs from the word in the utterance– for example, using a nickname like "bidiba" to consistently refer to a blanket would be coded as "blanket."

Conversely, non-concrete nouns can roughly be thought of as any noun that cannot be pointed to, such as abstract concepts. These nouns were not included. Gestures that could not be readily identified or that gestured to vague locations (e.g., "over there" while gesturing) also did not receive a referent code.

S1.2 Coding for Presentness

Not all concrete referents that are mentioned in speech are actually present in the environment. Using contextual information in the transcripts, we thus also coded whether each referent was likely to be present or not. Key contextual features included whether the same referent had been recently gestured to, question forms ("e.g.,"Where is the X¿'), and semantic information about presentness ("Look at that¿'). Every referent in an utterance is coded as"1" if present or "0" if not present. A subset of these codes were also done using video data to establish the reliability of our presentness coding. Further information can be found in the full coding manual at **XYZ.COM**

ADD FINAL LINK FOR FULL CODING MANUAL.

S1.3 Coding Example

person	chat	spoken obj	ref_pres_predicted
parent	want a bite of banana ?	banana	1
child	orange .	orange	0
parent	that's all the orange we have .	orange	0
parent	that's it .		

S2 Experimental Results

For readability, the main text includes only the key effects for each statistical model rather than a full specification. We include those here. Each model included at least a random intercept for each subject and item. Models were estimated using version XXXX of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015).

S2.1 Learning

To confirm that we sucessfully manipulated participants' learning, we asked whether items with more exposure during training were better learned at pretest. To do this we fit a logitistic mixed-effects model to analyze learning at baseline (i.e. prior to gameplay). We see the predicted signficant effect of exposure rate on learning, confirming that object-label mappings that were presented more in training were better learned.

Additionally, we can test our critical between-subjects manipulations to ensure that learning of the lexicon does not differ significantly at pretest (prior to the manipulations). Neither utility condition or partner's exposure signficantly predicts performance at pretest. This provides a simple check that participants in each condition learned the lexicon similarly. The full results of this model are presented in Table S1.

Table S1: Pariticpant learning at baseline, specified as `testCorrect ~ exposureRate + condition + (exposureRate | subj) + (1 | realLabel)`.

term	estimate	std.error	z-value	p-value
intercept	-0.97	0.23	-4.21	$< .001$
exposure rate	1.08	0.08	13.67	$< .001$
utility condition	-0.22	0.19	-1.15	.249
partner's exposure	0.11	0.12	0.93	.352

S2.2 Communicative Strategy

Our key analyses concerned participants choice of communicative strategy. In each trial, participants were able to choose one of 3 communicative strategies: gesture, speech, or teaching. We expected flexibile trade-off between the use of each strategy given their relative utilities, participant's knowledge of the lexicon, and the listener's knowledge of the lexicon. To test our predictions about each communicative behavior (gesture, speech, and teaching), we conducted separate logisitc

mixed effects models for each behavior, reported below. The mixed effects model for each communicative behavior has an identical effect structure for comprability. Each model includes a random effect of subject and item. It should be noted that these three behaviors are mutually exhaustive.

S2.2.1 Gesture

To examine gesture, we ran a mixed effects logistic regression to predict whether speakers chose to gesture during a given trial as a function of the target object's exposure rate during training, object instance in the game (first, second, or third), utility manipulation, and partner manipulation. Random effects terms for subject and object were included in the model. Consistent with our predictions, participants gestured more for words that recieved less training, with partners who had less knowledge, and in the condition where gesture's utility was higher. The full results of this model are presented in Table S2.

Table S2: Propensity to use gesture as a strategy, specified as `gesture ~ exposureRate*partnersExposure + appearanceNum*partnersExposure + utilityCondition + (1 | subj) + (1 | realLabel)`.

term	estimate	std.error	z-value	p-value
intercept	2.07	0.35	5.96	$< .001$
exposure rate	-0.50	0.04	-12.88	$< .001$
partner's exposure	-0.81	0.22	-3.61	$< .001$
instance	0.03	0.06	0.56	.576
lower speech efficiency condition	1.20	0.32	3.79	$< .001$
partner's exposure * exposure rate	-0.21	0.03	-6.69	$< .001$
partner's exposure * instance	0.07	0.04	1.53	.127

S2.2.2 Speech

To examine speech, we ran a mixed effects logistic regression to predict whether speakers chose to speak during a given trial as a function of the target object's exposure rate during training, object instance in the game (first, second, or third), utility manipulation, and partner manipulation. Random effects terms for subject and object were included in the model. Consistent with our predictions, participants used labels more for words that recieved more training, with partners who had more knowledge, and in the condition where speech's efficiency was higher. The full results of this model are presented in Table S3.

S2.2.3 Teaching

To examine teaching, we ran a mixed effects logistic regression to predict whether speakers chose to teach during a given trial as a function of the target object's exposure rate during training, object instance in the game (first, second, or third), utility manipulation, and partner manipulation. Random effects terms for subject and object were included in the model. Consistent with our predictions, participants taught labels more often (i.e. used both the gesture and speech strategy simultaneously) for words that recieved more training, with partners who had less knowledge, and

Table S3: Propensity to use labeling as a strategy, specified as `label ~ exposureRate*partnersExposure + appearanceNum*partnersExposure + utilityCondition + (1 | subj) + (1 | realLabel)`.

term	estimate	std.error	z-value	p-value
intercept	-3.19	0.30	-10.66	< .001
exposure rate	0.35	0.04	9.76	< .001
partner's exposure	1.95	0.19	10.02	< .001
instance	0.72	0.06	12.93	< .001
lower speech efficiency condition	-0.87	0.25	-3.42	.001
partner's exposure * exposure rate	0.27	0.03	9.05	< .001
partner's exposure * instance	-0.48	0.04	-11.07	< .001

Table S4: Propensity to use teaching as a strategy, specified as `teach ~ exposureRate*partnersExposure + appearanceNum*partnersExposure + utilityCondition + (1 | subj) + (1 | realLabel)`.

term	estimate	std.error	z-value	p-value
intercept	0.07	0.29	0.25	.799
exposure rate	0.14	0.04	3.21	.001
partner's exposure	-2.23	0.27	-8.38	< .001
instance	-1.09	0.07	-14.69	< .001
lower speech efficiency condition	-0.96	0.29	-3.29	.001
partner's exposure * exposure rate	-0.14	0.05	-2.70	.007
partner's exposure * instance	0.48	0.08	5.86	< .001

in the condition where speech's efficiency was higher. The full results of this model are presented in Table S4.

References