**Methods**

**Participants**

21 undergraduate students were recruited from the Stanford Psychology 1 credit pool. Of the participants, 11 were females and 10 were males. 20 were between the ages of 18 and 21, and one was below the age of 18. The students received course credit for participation. [sentence(s) about excluded participants] The final sample included [x] participants.

**Stimuli and Apparatus**

Eye-tracking software from SensoMotoric Instruments (SMI) was used to design the study and to collect the coordinates of participants’ eye gazes. The design of the video was based on a paradigm introduced by Yurovsky and Frank (2015) and further developed by MacDonald, Yurovsky, and Frank (under review) to include the presence of a social cue. The participants viewed the video on a 1920x1080 laptop screen.

The experiment featured sixteen pseudowords recorded by an AT&T Natural VoicesTM speech synthesizer using the “Crystal” voice (a woman’s voice with an American English accent), as well as 48 novel objects represented by black-and-white drawings of fictional objects from Kanwisher, Woods, Iacoboni, and Mazziotta (1997). Sixteen words were used so that the experiment would be sufficiently long to make within-subject comparisons across trials, and 48 objects were used so that objects would not be repeated across trials. The trial design is explained in the subsection immediately below.

**Design and Procedure**

Participants were seated with their faces about a foot away from a monitor and told they would watch a very short video, during which their eye movements would be recorded. They were asked to stay still and to keep their eyes on the screen. The experimenter then began the video, which lasted 2.6 minutes, and stepped away from the screen until the experiment was over.

The experiment consisted of a series of paired *exposure* and *test* trials, such that each exposure trial was immediately followed by a test trial. The first two pairs of trials were invariably *training* trials, followed by sixteen novel trials, whose order was randomized by the eye-tracking software, for a total of eighteen pairs of exposure and test trials. Furthermore, participants were randomly sorted into either the *gaze* or the *no-gaze* condition, explained in detail below.

On exposure trials, a woman’s face appeared on the screen above two novel objects, while a woman’s voice pronounced a pseudoword. In the *no-gaze* condition, the woman’s face looked straight ahead while the word was pronounced. In the *gaze* condition, the woman’s face turned toward one of the objects and then back to straight-ahead. The direction in which the woman’s face turned on a given trial in the gaze condition was random, but counterbalanced so that she turned towards the leftmost object eight times and the rightmost the other eight. A participant in the *gaze* condition was said to “follow gaze” if she spent a larger proportion of time on a given exposure trial looking at the object that was the target of gaze than at the other object.

On test trials, the woman’s face looked straight ahead in both conditions, while the voice repeated the same word from the immediately preceding exposure trial. One of the two objects from the exposure trial remained on the screen, while the other of the two objects was replaced by a third object that had not previously appeared in the video. To illustrate, if Objects A and B were displayed during an exposure trial, either Objects A and C or Objects B and C would be displayed during the following test trial. In what follows, the object that remained on the screen across both trials (in the illustrated case, Object A) is referred to as the “kept” object for convenience. A participant was said to “succeed” on a test trial if she spent a larger proportion of time looking at the kept object than at the new (non-kept) object.

[screenshots of exposure/test trial pairs in the gaze and no-gaze conditions]

The first two exposure/test trial pairs were *training* pairs. The objects in training trials were commonly recognizable objects, such as a squirrel or a cup, while the corresponding words were common English words that corresponded to an object on the screen, such as “squirrel” or “tomato”. The training trials were meant both to signal to the participant that the face was “labeling” objects on the screen and to check that participants were following the face’s gaze (in the *gaze* condition). Participants who did not reliably follow gaze on both training trials were excluded from the study. On the exposure training trials, the woman’s face looked straight ahead as the objects were labeled in the *no-gaze* condition, while in the *gaze* condition, she looked at one of the two objects, consistent with the later exposure trials. On the test training trials, the woman’s face looked straight ahead on both conditions.

The placements of the “kept” objects in both the exposure and test trials were counterbalanced throughout the experiment. Of the sixteen non-training test trials, eight “kept” the object that had been the leftmost of the two objects in the corresponding exposure trial. The other eight non-training test trials “kept” the object that had been on the right in the exposure trial. Additionally, the position of the “kept” object was counterbalanced such that it appeared on the left of the screen for half of the test trials and on the right for the other half of the test trials, independent of its position in the corresponding exposure trial. Finally, in the *gaze* condition, half of the test trials kept the object that was the target of the face’s gaze in the exposure trial, while the other half kept the object that had not been the target of gaze.

For the sake of analysis, the terms “Same trials” and “Switch trials” are used to dichotomize the test trials; the terms are used in a slightly different way than in MacDonald, Yurovsky, and Frank (2015), so a brief explanation is included here. “Same” refers to the test trials in which the kept object was also the object that the participant spent a larger proportion of time looking at during exposure; “Switch” refers to the test trials in which the object that was kept was the same object that the participant had looked at less during exposure. We hypothesized that participants in the *gaze* condition would allocate more attention and look more to the target of the face’s gaze on exposure; that they would thus be less likely to encode a link between the word and the non-target object; and that they would then perform well (that is, “succeed” often) on “Same” test trials and perform at chance (“succeed” 50% of the time) on “Switch” test trials. We predicted on the other hand that participants in the *no-gaze* condition would allocate equal attention to both objects on exposure and thus perform better than *gaze* condition participants on “Switch” trials.