Creating words from iterated vocal imitation

Edmiston et al. 10.1073/pnas.XXXXXXXXXX

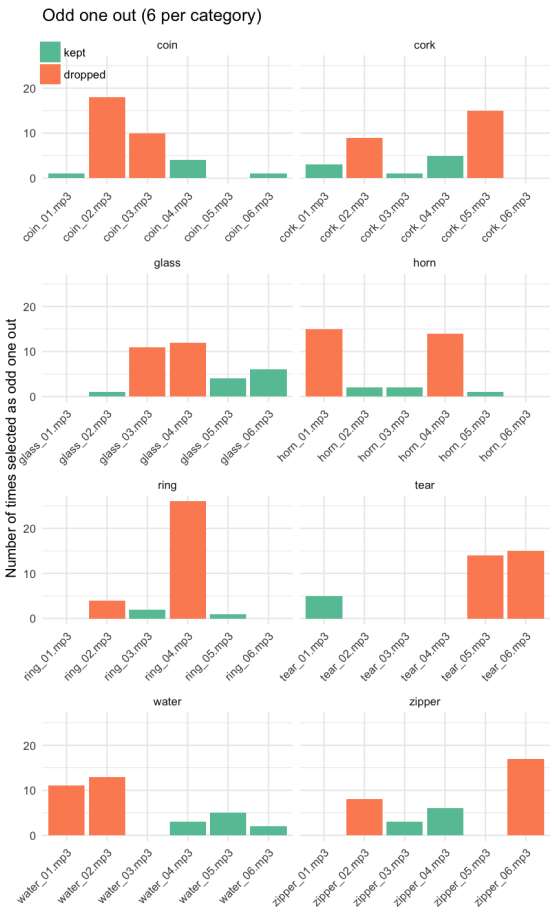


Fig. S1. Results of the first round of seed norming. After collecting these data, two sounds were removed from each category and the norming procedure was conducted again.

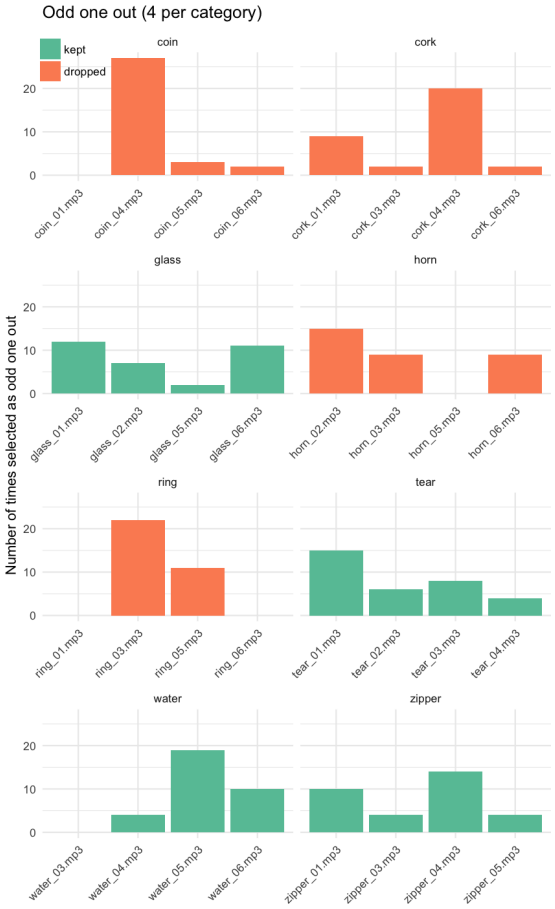


Fig. S2. Results of the second round of seed norming. After collecting these data, four categories of sounds were selected to use in the main experiment.

1. Selecting seed sounds

Our goal in selecting sounds to serve as seeds in the transmission chains was to pick multiple sounds within a few different categories such that each category member was approximately equally distinguishable from the other sounds within the same category. To do this, we started with an initial set of 6 categories and 6 sounds in each category and conducted 2 rounds of "odd one out" norming to reduce the initial set to a final set of 16 seed sounds: 4 sounds in each of 4 categories.

Participants in the odd one out norming procedure listened to the sounds in each category and picked the one that they thought was **the most different** from the others. In the first round of norming, participants listened to 6 sounds on a given trial. We removed the sounds in each category that were the most different from the others, and repeated the norming process again with 5 sounds in each category. The resulting sounds that were selected in each category are considered to be a set of equally distinguishable category members.

Final seed sounds.

Table S1. Environmental sounds used as "seed messages" in the experiment.

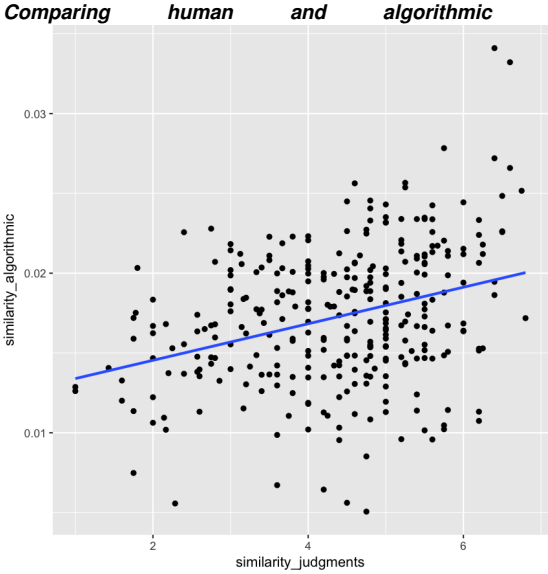
Category	Exemplar
glass	http://sapir.psych.wisc.edu/telephone/seeds/glass_01.mp3
glass	http://sapir.psych.wisc.edu/telephone/seeds/glass_06.mp3
glass	http://sapir.psych.wisc.edu/telephone/seeds/glass_02.mp3
glass	http://sapir.psych.wisc.edu/telephone/seeds/glass_05.mp3
tear	http://sapir.psych.wisc.edu/telephone/seeds/tear_01.mp3
tear	http://sapir.psych.wisc.edu/telephone/seeds/tear_03.mp3
tear	http://sapir.psych.wisc.edu/telephone/seeds/tear_02.mp3
tear	http://sapir.psych.wisc.edu/telephone/seeds/tear_04.mp3
water	http://sapir.psych.wisc.edu/telephone/seeds/water_05.mp3
water	http://sapir.psych.wisc.edu/telephone/seeds/water_06.mp3
water	http://sapir.psych.wisc.edu/telephone/seeds/water_04.mp3
water	http://sapir.psych.wisc.edu/telephone/seeds/water_03.mp3
zipper	http://sapir.psych.wisc.edu/telephone/seeds/zipper_04.mp3
zipper	http://sapir.psych.wisc.edu/telephone/seeds/zipper_01.mp3
zipper	http://sapir.psych.wisc.edu/telephone/seeds/zipper_03.mp3
zipper	http://sapir.psych.wisc.edu/telephone/seeds/zipper_05.mp3

Fig. S3. The interface for the telephone game. Initially the only action open to players is to hear the message by clicking the top sound icon. After listening to the message once, they could then initiate a recording of their imitation by clicking the bottom sound icon. Turning the recorder off submitted their response, and a new message was loaded.

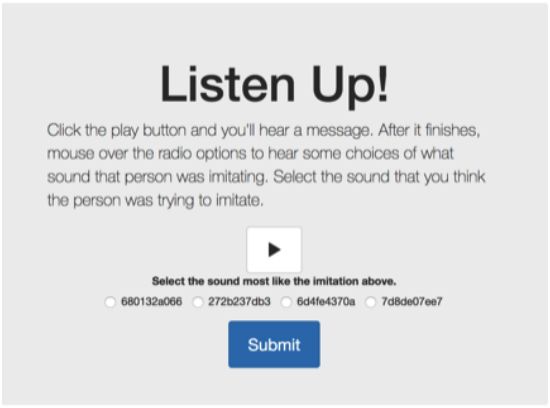
2. Collecting vocal imitations

Measuring acoustic similarity.

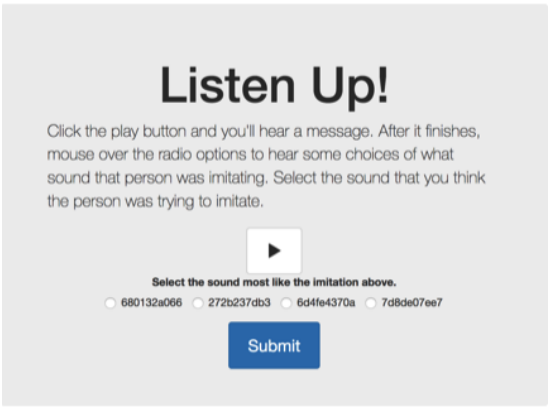
Comparing human and algorithmic similarities.



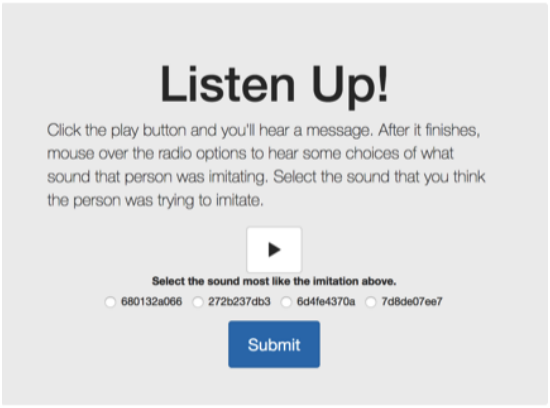
3. Matching imitations to seeds



4. Collecting transcriptions of imitations



5. Matching transcriptions to seeds



6. Learning transcriptions as category labels

To determine which transcriptions to test as category labels, we first selected only those transcriptions which had above chance matching performance when matching back to the original seeds. Then we excluded transcriptions that had less than two unique characters or were over 10 characters long, and sampled from both first and last generation imitations to reach a final set that controlled for overall matching accuracy.

Participants learned, through trial-and-error, the names for four different categories of sounds. On each trial participants listened to one of the 16 environmental sounds used as seeds

and then saw a novel word–a transcription of one of the imitations. Participants responded by pressing a green button if the label was the correct label and a red button otherwise. They received accuracy feedback after each trial.

The experiment was divided into blocks so that participants had repeated exposure to each sound and the novel labels multiple times within a block. At the start of a new block, participants received four new sounds from the same four categories (e.g., a new zipping sound, a new water-splash sound, etc.) that they had not heard before, and had to associate these sounds with the same novel labels from the previous blocks. The extent to which their performance declined at the start of each block serves as a measure of how well the label they associated with the sound worked as a label for the category.