Selecting words to learn as category labels in Experiment 3

Our procedure for sampling transcriptions to use as category labels was as follows: First, we removed transcriptions that contained less than 3 unique characters and transcriptions that were over 10 characters long. Of the remaining transcriptions, a sample of 56 were selected that were approximately equally associated with the target category.

To measure the association between each imitation and its target category (the category of the seed sound), we used the match accuracy scores reported in Experiment 2. The reason for using this measure of association strength as a control for selecting words to learn as category labels was to be able to select words that were initially equally associated with the target categories. Equating along this dimension allowed for a more focused test of differences between the words in terms of generalization to new category members.

The final sample of transcriptions were selected using a bootstrapping procedure which involved selecting a desired mean (the average association strength for eligible transcriptions of last generation imitations) and sampling transcriptions from first generation imitations and from seed sounds until the match accuracy of those imitations matched the desired mean within 1 standard deviation.