Online guessing experiments for iconicity measures

Bonnie McLean

2020-10-01

# Background

* Iconicity measures should become a standard part of word norms since they impact processing, but there still aren’t good measures for getting them.
* Recent review (Arthur) found problems with iconicity ratings from native speakers, unreliable. I should do a comparison with overlapping words.
* Mark’s thing Dingemanse et al. (2016) where guessability was correlated with their actual iconicity coding.
* Since corona etc. and crowdsourcing is very fast blah blah blah we provide a methodology for collecting iconicity ratings online using guessing experiments on the platform Prolific (talk about why you chose Prolific over AmazonTurk etc.). We have also made all the scripts and code for making the experiments freely available online and hope other people will use them.

# Method

## Materials

The recordings were synthesised using Google Cloud Text-to-Speech [(https://cloud.google.com/text-to-speech)](https://cloud.google.com/text-to-speech), with the female Japanese Wavenet voice ja-JP-Wavenet-B. They were then resynthesised using Praat to have a flat pitch of 200Hz. This was to eliminate the effect of intonation on the iconicity of the recordings.

## Participants

Participants were recruited through Prolific (www.prolific.co), and were self-identifying monolingual speakers of English residing in England (UK). Each participant participated in only one experiment. You rejected one participant because they failed the attention check, then after that you made it so they had to have an approval rating of 80% or more. A total of X participant’s responses were rejected because of failing attention checks, which meant the total number of responses was then XXX.

## Procedure

The experiments were conducted online, with each participant tested on 25 words. Participants were informed of the nature of the research before agreeing to participate, and were requested not to participate if they spoke or understood Japanese (see Figure 1). They were also instructed to wear headphones during the experiment, and could not start the experiment without first passing a listening test designed to require the use of headphones (Woods et al. 2017).

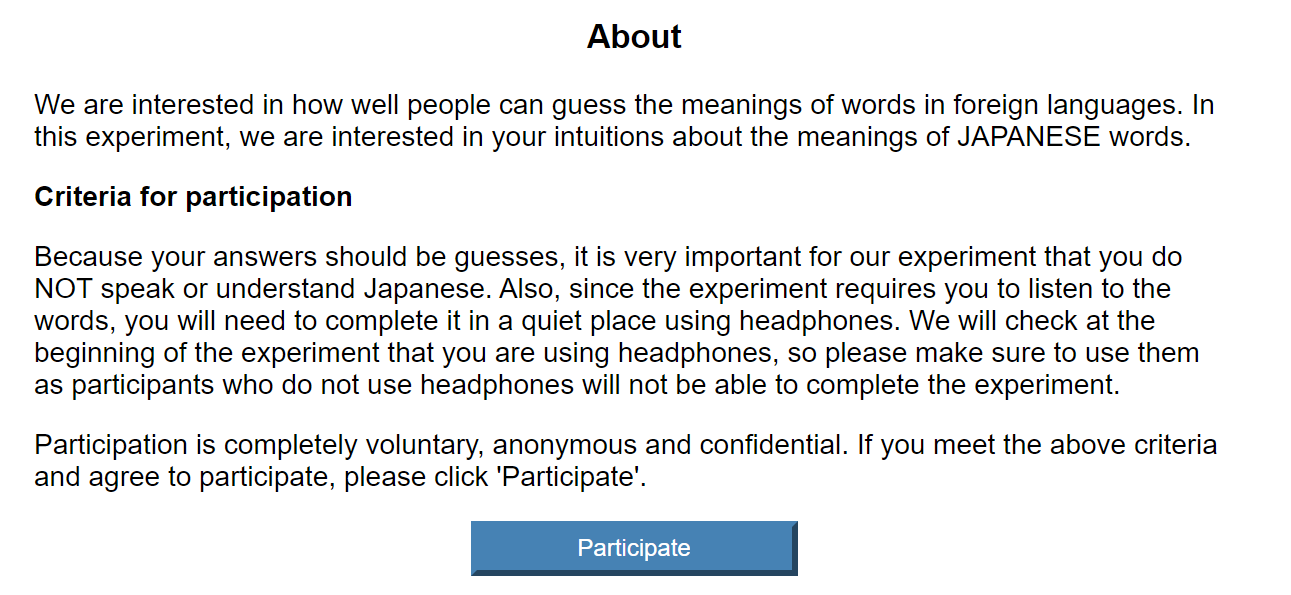


Figure 1: About page for experiment.

After passing the listening test, participants were taken to a page showing instructions for the experiment, and a button to start the experiment (Figure 2). Each experiment consisted of 25 trials. The first two trials in each experiment were fixed practice trials, the results of which are not included in the analysis. In addition, three trials were designed as ‘attention checks’, containing … COMPLETE LATER. These attention checks were randomly interspersed with the remaining 23 trials, which contained the forms whose iconicity was being tested.

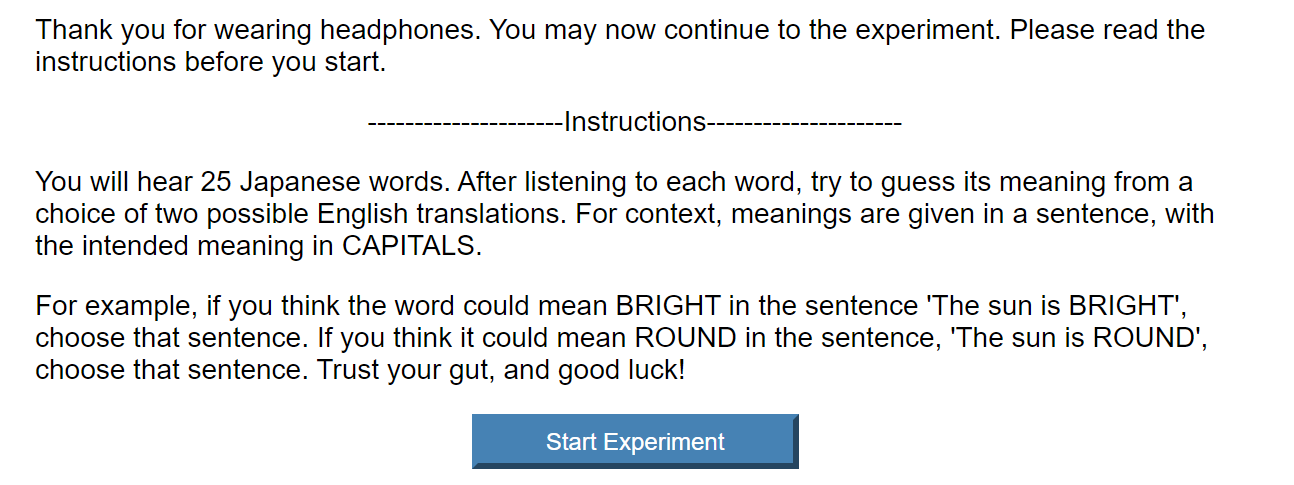


Figure 2: Instructions for experiment.

In each trial, participants were instructed to listen to a recording of the word. On playing the recording, they were shown two possible translations for the word, which they then had to choose between to continue to the next trial (see 3). Participants reaction times during the trials were also recorded, in order to screen out participants whose reaction times were significantly longer or shorter than average.

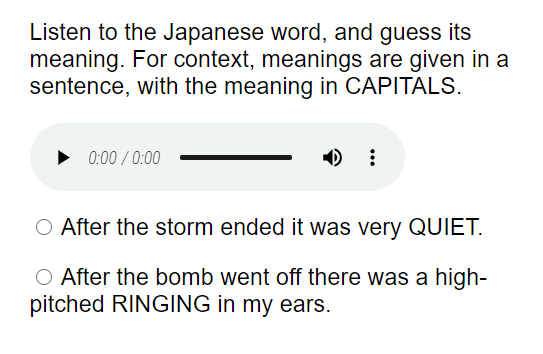


Figure 3: Sample trial in experiment.

Dingemanse, Mark, Will Schuerman, Eva Reinisch, Sylvia Tufvesson, and Holger Mitterer. 2016. “What Sound Symbolism Can and Cannot Do: Testing the Iconicity of Ideophones from Five Languages.” *Language* 92 (2): e117–e133.

Woods, Kevin J. P., Max Siegel, James Traer, and Josh H. McDermott. 2017. “Headphone Screening to Facilitate Web-Based Auditory Experiments.” *Attention, Perception & Psychophysics* 79 (7): 2064–72. <https://doi.org/10.3758/s13414-017-1361-2>.