Experimental Design

Activation of basic-level categories in the human brain based on different cue types

Background:

In this paper, we are replicating the experiment 1A from the paper "What makes words special? Words as unmotivated cues" (Pierce Edmiston, Gary Lupyan), in which the authors found out that spoken words activate conceptual knowledge about a category more efficiently than environmental sounds.

In our experiment, we seek to validate whether participants indeed react faster when being primed with a spoken word representing the basic-level category, e.g. the spoken word "dog" compared to when they are being primed with an environmental sound belonging to the basic-level category, e.g. any dog barking. The sound serves as an auditory representation of one token of the given category while the spoken word comprises the whole category. This difference coincides with the semantic embedding a word has, for example, usually people refer to a kind of furniture that you can sit on as a "chair" instead of saying the more general category "furniture" or the very specific one, e.g. "desk chair", "kitchen chair" etc.

If the spoken word, representing the entire category, is recognized faster than the environmental sound, it would mean that our mental concept of the basic-level category is activated more efficiently by the spoken word compared to the environmental sound.

Contrarily, if the environmental sound leads to a faster reaction time than the spoken word, one would need to further investigate this effect in another study.

Hypothesis:

We hypothesize that environmental sounds as auditory cues cause longer reaction times than spoken category labels when trying to assess and match their basic-level category to the image that is being displayed. We expect that environmental sounds activate a more specific category instance linking the sound to its likely source, although participants are instructed to only treat it as its basic-level category. This would mean environmental sounds are less effective cues than spoken category labels.

We also hypothesize that congruent[1] environmental sounds exhibit a faster reaction time than incongruent ones because there might be less mental effort involved in assigning congruent sounds to their corresponding basic-level category.

As exploratory hypotheses we want to find out whether the assumed effect from hypothesis one is still present if we exclude all incongruent trials or whether the hypothesized difference in reaction times is solely caused by them.

Furthermore, we want to explore whether the assumed effect from hypothesis one is also present when only considering mismatching trials.

---

[1] In our experiment, a trial is congruent if the environmental sound cue does not only match with the basic-level category but also with the specific instance of the image that matches the likely sound source.
Contrarily, incongruent means that the given environmental sound cue matches with the basic-level category but not with the specific instance of the category shown in the picture.

Participants:

Participants will mainly consist of family and friends, as well as fellow students. We try to test as many people as possible given the time window of one week. Our goal is to find at least 43 participants to reach the number of people that participated in the experiment in the paper.

Materials & Design:

We use 24 color photograph images that can be divided into 6 basic-level categories: *bird, car, dog, instrument, phone and typing*.

Each basic-level category comprises 4 images where two of them represent a different instance of the respective category (e.g. two images of a *Chihuahua* and two images of a *Rottweiler* for the basic level category "dog").

For each category we have one congruent and one incongruent environmental sound[2] and one spoken word for each basic-level category ('*bird', 'car', 'dog', 'instrument', 'phone', 'typing')* as auditory cue.

The materials can be found here:
https://github.com/jmdudek/XP-Lab2020-What-makes-words-special-Group-35/tree/master/Experiments/01_pilot/stimuli/sounds

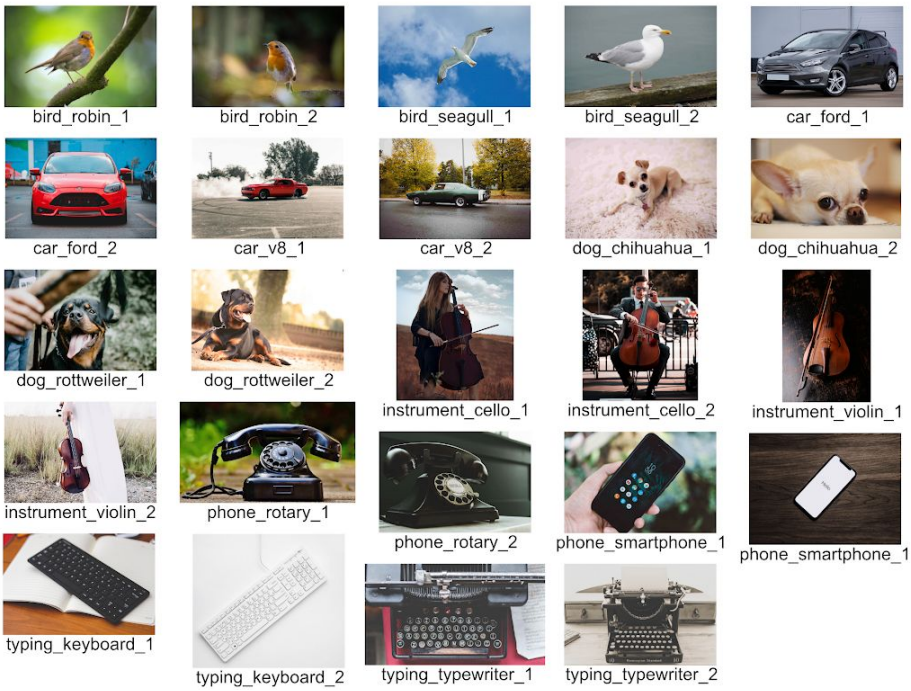
The peak amplitude of all sounds, spoken words and environmental sounds, is normalized to -10db and the duration of each environmental sound is set to 1s, using the free audio editing tool *audacity*.

We compressed the size of every image to 50kb by reducing the number of pixels while keeping its ratio, using photoshop.

To control for cue variability we use a female and male synthetic voice for each category name, to account for the bias sound pitches could possibly have on participant's responses.

[2] e.g. <Chihuahua bark> congruent for the Chihuahua images, incongruent for the Rottweiler images
<Rottweiler bark> congruent for the Rottweiler images, incongruent for the Chihuahua images
each representing the likely sound source for their respective instance of the basic-level category "dog"

Color photograph images that are used in the experiments:



Procedure:

In a trial, participants see a 250ms fixation cross in the middle of the screen after which they are primed immediately with a sound cue, i.e. a spoken word or an environmental sound.

After the offset of the sound, a target picture appears with a one second delay and is displayed until a decision is made by the participant.

Participants are instructed to minimize reaction time, while still answering as accurately as possible.

A decision is made by judging whether the auditory cue matches the picture that is displayed or not.

An auditory cue matches the picture if it belongs to the same basic-level category, i.e. any sound of a bird or the spoken word ‘“bird” followed by a picture of any bird.

There are congruent and incongruent environmental sounds for each basic-level category, which should both lead to a “yes” response.

The experiment starts with 6 practice trials in which the participants receive feedback via a popup window telling them whether their given answer was correct or not. In case of an incorrect answer, the popup window also gives a brief explanation why.

After the practice trials, participants are tested in 144 test trials, in which the cues match the pictures in 50% of the cases.

Participants indicate their decision by either pressing the "q" or "p" button on their keyboard. "q" indicates a yes response, for example, the sound of a cell phone ring or the spoken word "phone" followed by a picture of any phone (same basic-level category). Contrarily, "P" indicates a no response, for example, the sound of a cell phone ring or the spoken word "phone" followed by a dog (different basic-level category).

Trials are presented as a random sequence of factor combinations of cue type and congruence.

Study design

The experiment uses a 2 (cue type: label, sound) x 2 (match: same, different) x 2 (congruency: congruent, incongruent) factorial within-subjects design[3].

The dependent variable we are measuring is the reaction time (RT).

We are manipulating the cue type of each trial, the congruency and whether or not the auditory cue matched with the shown image.

Sources

Experiment 1A of "What makes words special? Words as unmotivated cues" (Pierce Edmiston, Gary Lupyan 2015).
Audacity: Audacity® software is copyright © 1999-2020 Audacity Team. The name Audacity® is a registered trademark of Dominic Mazzoni.
Sounds (all sounds were normalized and cut to 1s in length):

CC: "ringtone.wav" by davidferoli
CC: "5_cello_E2.wav" by flcellogrl
CC: "Aggressive Guard Dogs" by Oneirophile
CC: Bird Whistling, Robin, Single, 13.wav by InspectorJ

Synthetic voices: https://www.text2voice.org/
All the images that were we used stem from: https://www.pexels.com/

[3] all participants perform all trials