Design for an Experiment on the Facilitating Role of Language

Authors: Nels Acquistapace, Kyra Dunkel, Rosa Großmann, Johanna Linkemeyer

---

Background

In 2012, Lupyan and Thompson-Schill discovered a "label-advantage" where verbal cues activate conceptual knowledge about a category more effectively than nonverbal cues. When asked to rate whether an auditory cue matches the category represented in a picture, the researchers found that reaction times (RTs) were lower for verbal cues such as the word "dog" compared to nonverbal cues such as <dog bark>. Lupyan and Thomson Schill therefore suggest that language in the form of verbal cues allows for activation of conceptual representations in a special, more categorical way compared to nonverbal cues (Lupyan & Thompson-Schill, 2012).

In 2015, Edmiston and Lupyan built on these findings, arguing that verbal labels act as "unmotivated cues" that are context-independent and do not allow one to draw conclusions about specific features of a category member. For example, the label "dog" activates a general concept that includes all subcategories – in this context different breeds – of the category *dog* and therefore does not allow inference of specific features, such as the size of the dog. In contrast, nonverbal cues such as environmental sounds act as "motivated cues", i.e. upon hearing the cue, people infer a relationship of the cue to a specific instance of a category. That is, environmental sounds activate a representation of specific instances with concrete features of a category. For example, a low-pitched bark would likely activate representations of breeds with concrete features that match the sound of this bark, such as a big-sized dog – e.g. a Labrador. A schematic on the difference of motivated and unmotivated cues is provided in Figure 1. The researchers conducted a series of experiments to investigate whether environmental sounds are less effective than verbal labels in activating a general concept because they act as motivated cues, that rather activate a specific instance with concrete features of a category (Edmiston & Lupyan, 2015).

In the following, experiment 1a from the paper "What makes words special? Words as unmotivated cues" by Edmiston and Lupyan (2015) is reconstructed. While verbal labels and environmental

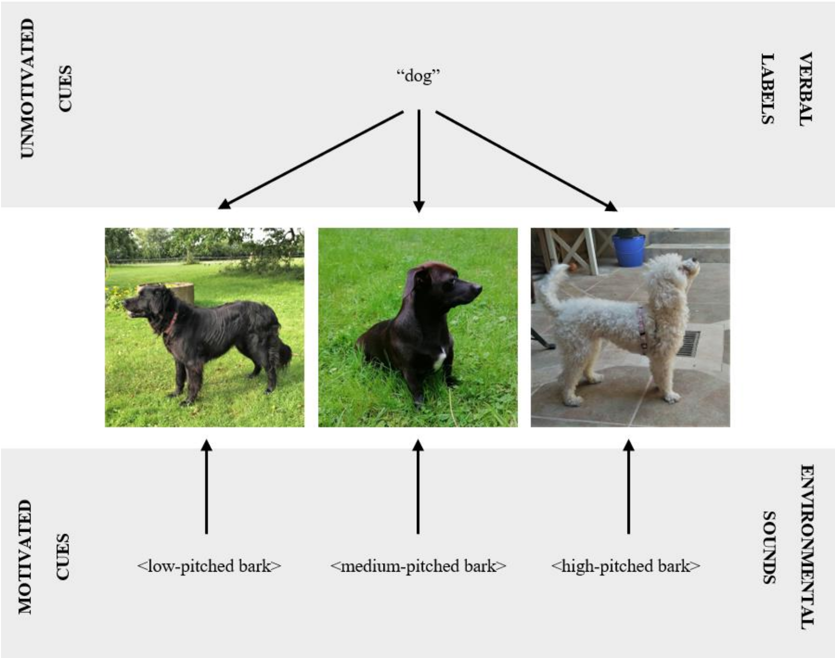

Figure 1. Examples of unmotivated and motivated cues for the category dog. The unmotivated cue (verbal label) "dog" refers to all three breeds (*Labrador mix, Dachshund mix and Miniature Poodle*). There is no specific pronunciation of the word "dog" in order to refer to a specific breed. The motivated cues (environmental sounds), the differently pitched barking, vary for each breed represented.

sounds can both be unambiguous representators for a specific category, this experiment tests whether environmental sounds activate likely sound sources and therefore fail to activate category-knowledge as effectively as verbal labels (Edmiston & Lupyan, 2015). For example, the label "dog" and the sound <dog bark> both refer to the category *dog*. However, the sound <dog bark>, as being a "motivated cue", would activate a more specific representation of the category *dog*, such as a dog of a specific size. To test the hypothesis, participants will hear an auditory cue that is either a verbal label or the sound of an object or animal (collaboratively referred to as environmental sounds). After a short delay, a picture is presented. Participants decide whether the previously heard auditory cue matches the category of the displayed object or animal. Trials starting with an environmental sound cue are divided into congruent and incongruent combinations. Congruent combinations are those where the environmental sounds match the subcategory presented in the picture. For example, hearing a low-pitched bark and seeing a Labrador. Incongruent combinations are those, where the basic-level category of auditory cue and presented picture is the same, but the stimuli represent different subcategories, as in hearing a low-pitched bark and seeing a Miniature Poodle. Importantly, participants are asked to decide merely based on the category-level and not on subcategory-levels of two stimuli. So, all congruent and incongruent combina-

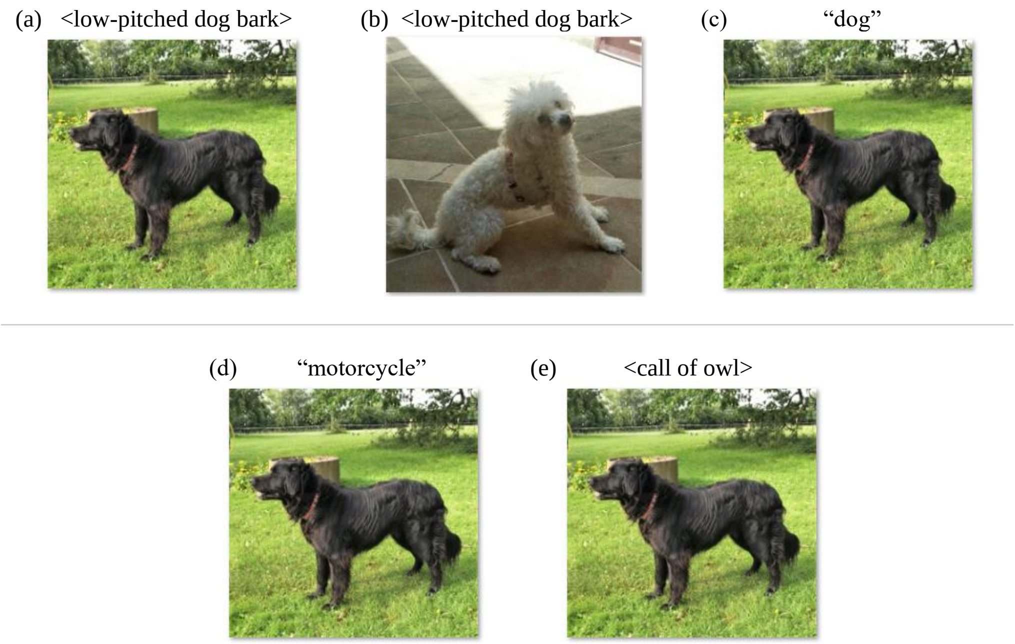

Figure 2. Different trial types for the picture category dog. Do these auditory cues match the category shown in the picture? Verbal labels are written in quotation marks, environmental sounds in angle brackets. The upper row shows trials where auditory cue and presented picture have matching categories. The row below shows trials where auditory cue and presented picture do not represent matching categories. (a) is an example for a congruent combination of stimuli. (b) represents an incongruent combination. (c) is an example for a matching label-picture combination. (d) shows an example of a nonmatching label-picture combination and (e) a nonmatching sound-picture combination.

tions are matching combinations and should be rated as such. A collection of possible trial types is shown in Figure 2. If participants perform better, i.e. they have faster RTs, for congruent trials than for incongruent trials this would allow for the conclusion that environmental sounds activate concepts of the sound source and not only the basic-level category.

Hypotheses

Through this replication experiment, different ways of activation of category knowledge through verbal labels and environmental sounds are investigated by addressing the following research hypotheses:

1. Sound cues are not as effective as category labels in activating a concept. That is, RTs for trials with sound cues are higher than for trials with verbal labels.
2. Sound cues activate specific category exemplars instead of general categories. That is, RTs for congruent trials will be lower than for incongruent trials.

To test these hypotheses, the RT needed for the participant to decide upon match/ nonmatch of the categories represented by auditory cue and picture is measured and is the dependent variable in this experiment. The experiment is a 2-by-2 factorial design; the independent variables are cue-type and sound-type. Each of the two variables has two levels. For cue-type, the levels are verbal label and environmental sound. Sound-type refers to the congruency of subcategories represented by the auditory and visual stimulus and therefore has two levels: congruent, and incongruent sound cue. All relevant manipulations of the independent variables are within-participants.

Low RTs indicate fast recognition of the stimuli categories – provided that participants answer as fast and accurate as possible. For the first hypothesis, the RTs from trial types with verbal cues will be compared to trial types with the environmental sound cues. To test the second hypothesis, RTs from congruent and incongruent sound cue trials will be compared.

Design

Materials. Materials for this replication study were obtained in different ways. All pictures except for one are provided by the authors, their friends, and acquaintances. Labels were recorded by the authors. Since the authors did not have access to the technical equipment to professionally record the environmental sounds, those were taken from a database. The feedback sounds used in this experiment (buzz and bleep) are widely known and associated with "incorrect" and "correct" feedback and were thus obtained from the materials of the original experiment. All materials, including German and English versions of a consent form participants have to agree with, are available for viewing-only purposes under this link.

There are four pictures for each of the six categories *bird, dog, guitar, drum, motorcycle, and phone,* creating a total of 24 pictures. Of each four pictures for one category, two correspond to one category, the other two to another subcategory. As an example, for the category *phone* there are two pictures of a cell phone and two pictures of a rotary phone provided. The subcategories chosen for the pictures are partly consistent with those of the original experiment. For those categories where different subcategories were selected (categories *bird, dog,* and *motorcycle*), subcategories were chosen according to the style of selection of the original experiment. All pictures are provided in square format and are shown in Figure 3.

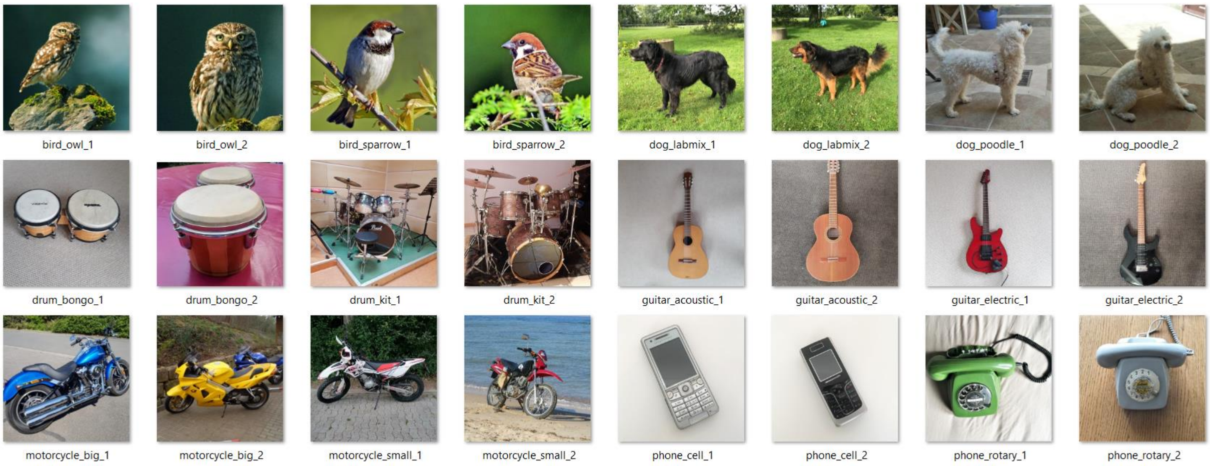

Figure 3. Collection of all pictures used in this experiment.

The 24 auditory cues used in this experiment consist of 12 verbal labels and 12 environmental sounds. For each of the six categories listed above, there are four auditory cues in total. Those four consist of two verbal labels, one spoken by a male and one by a female. The other two auditory cues are environmental sounds, chosen for each subcategory of the six categories. For example, for the category dog the subcategories are *Labrador mix* and *Miniature Poodle*. So, the environmental sounds for the category dog are <low-pitched bark> for the subcategory *Labrador mix* and <high-pitched bark> for the subcategory *Miniature Poodle*. All auditory cues were equalised in length (600ms) and normalized in volume.

Main Trials. In total, the main experiment consists of 192 trials with a 50:50 division to matching and nonmatching combinations. There are 96 possible matching combinations, i.e. trials where the auditory cue category matches the picture's category, of the 24 pictures and 24 sounds described above. Nonmatching combinations of the categories represented by the auditory cue and the picture are included to increase participants' caution on the task during the experiment. The nonmatching trials are constructed by the authors by first creating all possible nonmatching label-picture and sound-picture combinations, which are 240 each. Of those combinations, 48 label-picture and 48 sound-picture combinations are chosen randomly, creating a total of 96 different nonmatching combinations. Each combination of the matching and nonmatching ones was tested twice in the original experiment. To avoid the order effect of fatigue or boredom through the repetitive style and length of the original experiment, each combination will only be tested once in this replication experiment.

Practice Trials. There are six practice trials of which three consist of matching and three of nonmatching combinations. These combinations are chosen from all combinations used in the main trials. As the practice trials are supposed to help the participants to better understand the task, the trials are carefully selected and contain all five trial types explained in Figure 2. This type of trial selection allows participants to familiarize themselves with all different trial types before starting the main experiment. Especially the inclusion of congruent and incongruent trials in the practice is important. Participants have to understand that the decision they have to make is solely based on the basic-level category of sounds and objects and not on subcategories, i.e. congruent and incongruent combinations both should be rated as matches. For example, <low-pitched bark> and a picture of a *Labrador mix* make a congruent combination. The sound <high-pitched bark> and a picture of a *Labrador mix* make an incongruent combination. Though differing in congruency, both trial types create matching combinations and should be rated as such.

Procedure. The experiment consists of five parts.

I. Introduction and instructions
II. Practice phase
III. Further instructions
IV. Main test phase
V. Post-experiment questionnaire

First, participants are welcomed to the experiment. They receive instructions on the experimental setup. The instructions are provided in English and German. We try to recruit as many participants as possible and we do not want to restrict the experiment to fluent English-speakers since a moderate English-level suffices for the purpose of this experiment. The instructions given are described below.

(1) Participants are informed that there is a consent form they automatically agree with by continuing the experiment. The consent from is available in English and German and can be accessed via a link.

(2) Participants are advised that their full auditory and visual attention is required. They are asked to run the experiment on a laptop or computer and use headphones during the experiment.

(3) Participants get information on the trial construct. That is, each trial starts with an auditory cue, followed by a picture that is presented after a short delay phase. Participants receive explanation of their task and further information on incongruent trials. Those trials might be counterintuitive to judge, as the subcategories of the stimuli do not match. Participants are informed that incongruent trials belong to the group of matching combinations.

(4) Next, participants receive information on the keys they have to use in the experiment. “y” is the key to be pressed for matching combinations and “n” should be pressed for nonmatching combinations. In this view, participants are also advised to adjust their volume during the practice trials. This is important because participants will be asked not to change the volume during the main trials.

During the practice phase, participants will get accustomed to the task by completing the six practice trials described above. The six sound-picture combinations that were chosen by the authors are shown in a random order, randomised for each participant. After the practice trials, participants are asked to press a button to start the main trials and optimise speed and accuracy of their answers. For both the practice and the main test trials, participants receive immediate feedback as to whether their response was “correct” or “incorrect” through a buzz or bleep sound. In the practice trials, participants also receive visual feedback in form of the words “correct” and “incorrect” to familiarize themselves with the feedback sounds.

Each trial of the practice and main task is structured the following way:

(1) Each trial starts with 500ms blank screen and no sound.

(2) A fixation cross appears centrally for 250ms.

(3) Then, an auditory cue is given (600ms).

(4) Next, there is a pause of 1s after the offset of the auditory cue.

(5) After the pause, a picture appears centrally on the screen. The picture size is automatically adapted to the screen size. The picture remains visible until the participant responds by pressing one of the two keys “y” or “n”, representing “yes” and “no”. Participants are asked to answer “yes” for matching combinations (e.g. <call of owl> or “owl” followed by a picture of an owl) and “no” for nonmatching combinations (e.g. <cell-phone ring> or “phone” followed by a picture of bongos). There is no time limit for the answer. Participants receive immediate feedback on the correctness of their answer. The order of the presented pairs of stimuli is randomised (shuffled on the fly for each participant).

(6) When finished with the main trials, participants can voluntarily provide socio-demographic information and feedback in a post-test survey.

(7) Finally, the last view thanks the participants for their contribution and asks them to press a button in order to submit their results.

Note on Replication Limitations

This experiment is an online experiment and is not consulted in a lab as the original experiment. Therefore, some modifications of the experiment were necessary. First, instead of being asked to answer using buttons (“yes”, “no”) on a controller as in the original experiment, this replication experiment is a forced-choice key-press experiment. This way, the experiment is easily executable at home on devices with a keyboard, such as laptops and computers.

In the original experiment, pictures subtended ~10x10° when participants were sitting approximately 24” (about 60cm) away from the screen. To include this relation of picture size and distance from the screen in this replication experiment, information on the pixel size of the device the experiment is conducted on would be required. As this is an online experiment, this requirement cannot be fulfilled.

Therefore, neither picture size nor participants' distance to the screen is fixed in this replication study. Rather, picture size will be automatically adapted to the participants' screen size and participants are instructed to sit away from their screen such that they are able to fully see the screen.

Concerning the volume of the auditory cues, participants are asked to adjust the volume during the practice trials and not to change the volume during the main trials. There is no information whether the volume was controlled between participants in the original experiment. By asking participants not to change their volume during the main trials, within-subject artifacts are tried to be avoided. Nevertheless, there is no possibility to control for this requirement. So, results might be altered in an unpredictable way.

Another note on limitations of this experimental design compared to the original experimental design is that technical differences (e.g. different keyboards and screen-types) of the participants' devices might alter results.

Shortly after the study was online and some participants took part, technical problems were detected. We were able to quickly link these problems to the Safari browser. Therefore, we added another instruction in retrospect, asking the participants to use a different browser than Safari when participating in this experiment.

References

Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Elsevier*, pp. 93-100.

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, pp. 170-186.