

# Methodological Details - Blue Bananas

9/17/2020

## 1. Research Questions

This study investigates whether speakers modify the speech signal to signal atypical referents? Concretely, we ask whether the suprasegmental profile of an utterance with an atypical referent like “blue banana” is different from one with a typical referent such as “yellow banana”.

## 2. Participants

30 native German speakers participated in this study. All participants grew up in a monolingual environment and were recruited from the population living in the broad Cologne area with normal or corrected-to-normal vision and normal hearing. Participants were paid for their participation.

## 3. Procedure

In this production study, participants interacted with the experimenter. Participants had to verbally instruct the experimenter to select a specified target object out of four visually presented objects. The non-target objects differed from the target with respect to their colour, their identity, or both. Objects were referred to using noun phrases that consist of a modifier denoting colour and a modified object (e.g. yellow banana, blue banana). These adjective-noun combinations differed with respect to the typicality of their combination. The adjective-noun combinations were either typical, medium typical, or atypical (as established in a norming study, see additional materials). Some combinations were typical such as yellow banana, red tomato, or yellow apricot; some were atypical such as red banana, purple tomato, or blue apricot; and some were somewhat in the middle like brown banana, green tomato, and red apricot (see )

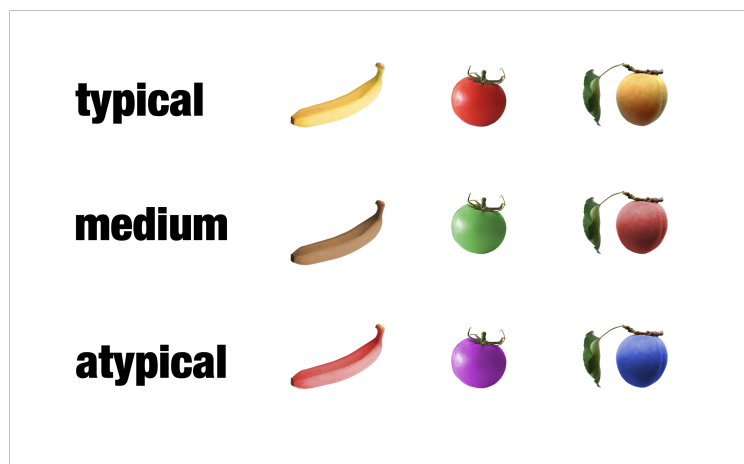


Figure 1: Three examples for typical, medium typical and atypical combinations of colour and object.

Participants were seated in front of a computer screen. The experimenter sat at the opposite side of the table in front of another computer screen. Participants and experimenter could see neither each other nor each others' screens. The experiment consisted of two phases. A familiarization phase and a test phase.

In a familiarization phase, participants saw one object per trial. In order to advance to the next trial, participants had to read out loud the corresponding noun phrase (e.g. blue banana). During this phase, participants had to name all atypical colour-object targets that we used in the test phase of the experiment alongside their typical counterparts. For example, if red banana (atypical) was an experimental target, participants were presented with both the red banana (atypical) and the yellow banana (typical). This familiarization phase was included in order to ensure that participants can relate typical and atypical colour-object combinations to each other.

After the familiarization phase, participants entered the test phase. On each trial in the test phase, participants first saw four coloured objects in the top left, top right, bottom left, and bottom right of the screen, respectively. One of the object served as the target; another served as the competitor; and the remaining two served as unrelated distractors. The position of the visual stimuli was randomized for each trial and each participant. In the center of the screen, a black cube was displayed which could be moved by the experimenter. The participants were asked to instruct the experimenter to move the cube onto one of the four pictures.

Each test trial consisted of two parts. The 'trigger' instruction and the 'test' instruction. After the preview of all images was displayed for 1500 ms, a 'trigger' sentence containing the competitor was orthographically presented below the cube:

"Du sollst den Würfel auf der grünen Sonnenbrille ablegen." engl. "You have to put the cube on top of the green sunglasses."

The trigger sentence was supposed to set a discourse context, such that the colour and the competitor referent were introduced as background information (here: green and sunglasses). The competitor referent was visually indicated by an arrow. The trial proceeds when the experimenter has moved the cube onto the respective referent. Subsequently, the sentence and the arrow disappeared and participants were presented with a new sentence containing the target referent:

"Und jetzt sollst du den Würfel auf der grünen Avokado ablegen." "And now, you have to put the cube on top of the green avocado."

Again, the target referent is visually indicated by an arrow. The sentence is orthographically displayed in the middle of the screen. When the experimenter completed the move of the cube to the respective referent, the trial was completed and the next trial initiated. There was a 3000 ms inter stimulus interval (grey screen) between trials. In both competitor and target sub sequences, the arrow was displayed with a lag of 1000 ms and the target sentence with a lag of 1500 ms in order to give participants sufficient time to glance at all referents before they see the arrow and the target sentence.

Using the trigger sentence to set up the discourse context enabled us to manipulate the focus structure of the target sentence. If the colour of the competitor and the target are of the same object type but differed in colour, the colour adjective was in focus (henceforth the Adjective Focus (AF) condition). If the objects differed but not their colour, the noun was in focus (henceforth the Noun Focus (NF) condition). If both the colour and the object differed, the whole noun phrase was in focus (henceforth the Adjective Noun Focus (ANF) condition).

While we included all three focus conditions, the study was originally planned to only look at Noun focus. AF and ANF serve as filler trials. There are 15 NF trials, and 10 AF and ANF trials each. Each trial occurs two times, yielding a total of 70 trials.

The code of the experiment is available [here](#)



Figure 2: Example screen at the beginning of a trials. The discourse setting trigger sentence is displayed in the middle below the cube. The participants have to instruct the experimenter to move the cube on top of the green sunglasses (indicated by an arrow).



Figure 3: Example screen after the trigger sentence. The target sentence is displayed in the middle of the screen. Now, participants have to instruct the experimenter to move the cube on top of the green avocado (indicated by an arrow).

## 4. Stimuli

### 4.1 Norming study

Images of colored referents serve as visual stimuli. The selection of stimuli was based on a norming study, as described in the following.

We selected twenty fruits or vegetables (henceforth: FOODs) and four other referents (henceforth: NON-FOODs). Selection criteria were mainly informed by visual discriminability and phonemic composition, such that referential expressions were supposed to contain many voiced segments. We created different visual versions of each individual referent. Each version was created by manipulating a photograph of the object in the image editor gimp. The photographs were collected from internet databases containing copyright-free high-resolution images (e.g. Pixabay). During the manipulation process, the background of the pictures was replaced by a white background. In order to change the colour of the objects, a layer in the respective colour was created and overlaid on top of the original graphics using “Colour” as the layer mode. The hexadecimal colour codes for the colours used are listed here:

Blue: #29429f Green: #3f8535 Red: #9d1c1c Purple: #6d1b79 Brown: #2a1d11 Yellow: #e3c917 Orange: #ff8400 Orange (potatoes): #ff6600 Black: #000000 Grey: #FFFFFF

The same colour code was used across referents to achieve a particular colour, except for orange, which had to be adjusted for potatoes to be more distinct from the original colour brown. To achieve grey colouring, a white layer was used – the interaction between the original layer and a white overlay actually resulted in a grey colour on the object. To achieve a darker black on certain fruits, the colour spectrum of the original picture had to be manipulated using Gimp’s Colour Curves tool. This tool allowed us to darken or brighten certain colour channels of the picture. Parts of the fruits that did not have the body colour of the fruit (e.g. like stalks and leaves) were left in their original colour. All pictures with natural and manipulated colors are available [here](#).

The colours for the twenty FOODs were chosen as to be perceived as having different degrees of typicality for each respective object, e.g. a banana is typically yellow, a less typical banana would be brown or green and an atypical exemplar could be blue. Every FOOD was presented in four of nine manipulated colours alongside its original colour. We assumed that the four NON-FOODs did not have a prototypical colour. We thus expected different colours to be comparably typical for these objects. The German nouns for all objects were either feminine singular or they occur in their plural form (e.g. grapes).

We used the following colours: black, blue, brown, green, grey, orange, purple, red, yellow.

We used the following FOODs: apricot, avocado, banana, beans, carrot, cherry, cucumber, eggplant, grapes, lemon, mandarine, pear, pepper, pineapple, potatoes, strawberry, tomato, walnut, zucchini.

We used the following NON-FOODs: clothespin, paper clip, socks, sunglasses.

100 German native speakers participated in the norming study using the crowd sourcing platform [Prolific](<https://www.prolific.ac>). We presented all objects in different colours to our subjects. FOODs came in 5 colours per object, NON-FOODs came in all nine colours. Subjects were instructed to rate how typical they think the colour for each object was, using a smooth slider ranging from 1 to 100 (see javascript for the experimental design [\[here\]](#))(<https://github.com/stelaseldano/colour-typicality-norming>).

### 4.2 Selection of experimental stimuli

The norming ratings were subsequently used to choose appropriate stimuli for the production study, using the following procedure. The results of the norming study can be retrieved [here](#).

First, we chose five colours from the norming data set: yellow, green, red, orange, and brown. Those were the most frequent colours in the superset and their norming ratings varied strongly as a function of the referent.

Second, we sorted the FOODs according to their typicality ratings and binned them into typical, atypical, and medium typical. Typical FOODs are defined by norming ratings above 90. Atypical FOODs are defined by norming ratings below 25. Mediocre typical FOODs are defined by norming ratings in between 25 and 90. For each colour and typicality, we selected one FOOD. Each cell was occupied by a different FOOD.

For the critical focus condition (NF) we thus had 15 target FOODs (5 colours x 3 typicality categories), as given below:

atypical: Yellow cherry (Gelbe Kirsche) Green carrot (Grüne Möhre) Red cucumber (Rote Gurke) Orange grapes (Orangene Trauben) Brown pepper (Braune Paprika)

medium: Yellow peas (Gelbe Erbsen) Green tomato (Grüne Tomate) Red apricot (Rote Aprikose) Orange potatoes (Orangene Kartoffeln) Brown banana (Braune Banane)

typical: Yellow lemon (Gelbe Zitrone) Green green beans (Grüne Bohnen) Red strawberry (Rote Erdbeere) Orange mandarine (Orangene Mandarine) Brown walnut (Braune Walnuss)

The set of 15 competitors from the NON-FOOD subset is chosen such that there are as many distinct competitors in each colour as there are targets. The set consists of:

Yellow sunglasses (Gelbe Sonnenbrille) Yellow socks (Gelbe Socken) Yellow clothes peg (Gelbe Wäscheklammer) Green sunglasses (Grüne Sonnenbrille) Green socks (Grüne Socken) Green Paper clip (Grüne Büroklammer) Red socks (Rote Socken) Red paper clip (Rote Büroklammer) Red clothes peg (Rote Wäscheklammer) Orange socks (Orangene Socken) Orange paper clip (Orangene Büroklammer) Orange clothes peg (Orangene Wäscheklammer) Brown sunglasses (Braune Sonnenbrille) Brown paper clip (Braune Büroklammer) Brown clothes peg (Braune Wäscheklammer)

The pairing of targets and competitors was randomized for each subject, with the only constraint that the competitor was of the same colour as the target referent but a different object (e.g. if blue banana was the target, blue grapes was the competitor) (see randomization section below for more detail).

For the filler focus conditions (AF + NAF), we distributed FOODs and NON-FOODs, such that over the whole experiment (a) the frequency of target FOODs is counterbalanced (3 times each) and (b) the frequency of target colour is counterbalanced (14 times each).

For all three conditions, a subset of visual distractors was chosen out of our superset from the norming study: Distractors are FOODs that are not used as targets or competitors (avocado, egg plant, pear, zucchini) and all NON-FOODs. They could have any of the five colours. Within each given trial, the distractors did neither share colour or object identity with the target or the competitor.

Overall, the sets of target, competitors and distractors are chosen such that all five colours occur equally often throughout the experiment, i.e. 28 times.

### 4.3 Randomization

For each trial, targets, competitors, and distractors are chosen in a randomized fashion, fulfilling the following requirements:

- NF-Condition: Colour of Target = Colour of Competitor Object of Target != Object of Competitor
- AF-Condition: Colour of Target != Colour of Competitor Object of Target = Object of Competitor
- ANF-Condition: Colour of Target != Colour of Competitor Object of Target != Object of Competitor
- All conditions: Both distractors have to be different in colour and object from target and competitor.

The competitor and distractors are matched iteratively to a target. One competitor can only be used twice or a third time if all other competitors that are still left have been considered and prove inappropriate.

The quadruples for the three conditions are formed separately. After the three lists have been compiled, they are merged. The sequence of quadruples is randomized such that the following requirements are met:

- The target colour of one quadruple must not equal the competitor colour of the next to avoid a contrastive focus on the next competitor noun
- The target object of one quadruple must not equal the competitor object of the next to avoid a contrastive focus on the next competitor adjective
- The list can have the same experimental condition in adjacent trials only in 12% of cases at the maximum

One randomised list for each subject is produced and doubled, so that each unique trial occurs twice during a recording session. In case the doubling of the list leads to violation of the above criteria, the first trial of the repetition list is moved until the criteria are met.

The python code for the randomization is available [here](#).