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

Thirty 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.

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. The participants and the 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, the competitor object was visually highlighted by an arrow, and a

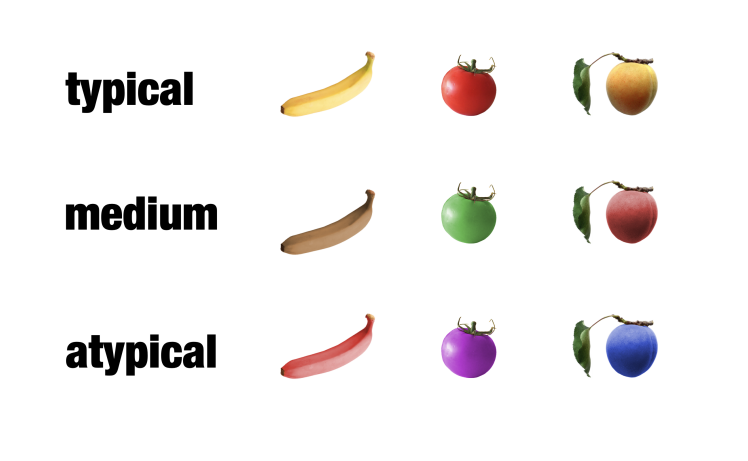

Figure 1: Three examples for typical, medium typical and atypical combinations of colour and object.

'trigger' instruction was orthographically presented below the cube (see Figure #):

(1) *Du sollst den Würfel auf der grünen Sonnenbrille ablegen.*

du sollst den würfel auf der grünen sonnenbrille ablegen.
you have.to the cube on the green sunglasses put

'You have to put the cube on top of the green sunglasses.'

The trigger instruction was supposed to create a discourse context, such that both the color and the competitor object were introduced as background information into the discourse (here: green and sunglasses). The trial proceeded when the experimenter has moved the cube onto the respective referent. The sentence and the arrow disappeared. Subsequently, the target object was visually highlighted by an arrow and the test instruction was presented containing the target object (see Figure #):

(2) *Und jetzt sollst du den Würfel auf der grünen Avokado ablegen.*

und jetzt sollst du den würfel auf der grünen avokado ablegen
and now have.to you the cube on the green avocado put

'And now, you have to put the cube on top of the green avocado.'

The trial was completed and the next trial initiated as soon as the experimenter had moved the cube to the target referent. There was a 3000 ms inter stimulus interval between trials (grey screen). 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 were able to identify the relevant object and the instructions.

Using the trigger instruction to set up the discourse context enabled us to manipulate the focus structure of the target sentence. If the color of the competitor and the target are of the same object type but differed in color (e.g. yellow banana –> blue banana), the color adjective was discourse-pragmatically focused (henceforth the Adjective Focus (AF) condition). If the objects differed but not their color (e.g. yellow banana –> yellow tomato), the noun was in focus (henceforth the Noun Focus (NF) condition). If both the color and the object

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).

differed (e.g. yellow banana –> blue tomato), the whole noun phrase was in focus (henceforth the Double Focus (DF) condition).

We are interested in the Noun Focus condition, only. AF and DF served as filler trials. There were 15 NF trials, 10 AF and 10 DF trials each. Each trial occurred two times per participants, yielding a total of 70 trials per participant.

The experimental program (implemented as a browser-based application) is available here: https://github.com/SBRitter/fruits-production.

A Additional Information

Here we specify in detail, how stimuli were selected for the above described production study.

A.1 Selection of stimuli superset

The selection of stimuli was informed by a norming study for which 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 nouns denoting our objects were supposed to contain few voiceless segments.

For each of the FOODs and NON-FOODs, we created colored versions by image manipulation using gimp. Images of all individual objects 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 color of the objects, a layer in the respective color was created and overlaid on top of the original image using "Color" as the layer mode. The hexadecimal color codes for the colors that we used are listed below:

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

All pictures with natural and manipulated colors are available here: https://osf.io/rdtx5/.

The colors for the twenty FOODs were subjectively selected to reflect a range of compatibilities with the respective objects, e.g. a typical banana is yellow; a less typical banana could be brown (too ripe) or green (not yet ripe); and an atypical banana could be blue. Every FOOD was presented in four of nine manipulated colors alongside its original image. The four NON-FOODs were subjectively selected to represent color agnostic objects which can and do come in a variety of colors. All objects were chosen such that their respective German nouns were either feminine singular (e.g. "die Banana" 'the banana') or plural (e.g. "die Trauben" 'the grapes').

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.

A.2 The norming study

One-hundred German native speakers participated in a norming study using the crowd sourcing platform [Prolific]](https://www.prolific.ac). We presented all objects in different colors to our participants. FOODs

came in 5 colors per object, NON-FOODs came in all nine colors. Participants were instructed to rate how typical they think the color 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. The results of the norming study can be retrieved here: https://osf.io/znpg5/.

A.3 Selection of experimental stimuli

The norming ratings were subsequently used to select appropriate stimuli for the production study according to the following procedure:

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

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

For the critical focus condition (NF) this selection procedure resulted in 15 target FOODs (5 colors x 3 typicality categories):

atypical:
Yellow cherry (Gelbe Kirsche)
Green carrot (Grüne Möhre)
Red cucumber (Rote Gurke)
Orange grapes (Oangene 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 was selected to ensure that there are as many distinct competitors for each color as there are target objects. The set consisted of the following objects:

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 had the same color as the target object (e.g. if blue banana was the target, blue grapes is a possible competitor) (see randomization section below for more details).

For the filler focus conditions (AF + DF), we distributed FOODs and NON-FOODs to fulfil the following constraints: (a) each target FOOD appeared 3 times and (b) each color appeared as a traget color 14 times.

For all three focus conditions, a subset of distractors was selected from the initial superset: Distractors were either FOODS that were neither used as targets nor as competitors (avocado, egg plant, pear, zucchini) or NON-FOODs. Within each given trial, the distractors did neither share color nor object identity with neither the target nor the competitor.

Overall, the sets of target, competitors and distractors are combined such that all five colors occurred equally often throughout the experiment (= 28 times).

B Randomization

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

- NF-Condition:
  Color of Target = Color of Competitor
  Object of Target != Object of Competitor
- AF-Condition:
  Color of Target != Color of Competitor
  Object of Target = Object of Competitor
- DF-Condition:
  Color of Target != Color of Competitor
  Object of Target != Object of Competitor
- All conditions:
  Both distractors have to differ in color and object identity from both target and competitor.

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

The set of four objects (quadruple) was formed separately for each condition. After the three lists had been compiled, they were merged. The sequence of quadruples was randomized such that the following requirements were met:

- The target color of one quadruple must not be equal to the competitor color of the following quadruple in order to avoid a contrastive focus on the next competitor noun.
- The target object of one quadruple must not be equal to the competitor object of the following quadruple to avoid a contrastive focus on the next competitor adjective.
- The list can have the same experimental condition in adjacent trials only in maximally 12% of all cases.

Following these criteria, one randomized list for each subject was produced. The list was then copied and merged with the first list in order to ensure that each unique trial occurs twice during one recording session. In case the copy-and-merge procedure led to violation of any of the criteria described above, the first trial of the repetition list was shuffled until the criteria were met.

The python code for the randomization is available here: https://github.com/SBRitter/fruits-randomization.