

Supplementary Online Material (SOM) to the Article:

Do People Remember What is Prototypical?

The Role of Accent-Religion Intersectionality for Individual and Category Memory

Appendix A

Material Selection

In order to select appropriate stimuli, we pretested a sample of preselected voices and faces ($N = 24$). Twenty-nine participants (8 men, 21 women, $M_{\text{age}} = 24.72$, $SD = 6.76$) from a German university evaluated the entire sample of female photographs and voice recordings in a randomised order. Participants used in the pre-test were drawn from the same population as those later taking part in the main experiment though they did not participate in both.

The choice to use only female targets came naturally as in this case we were able to indicate Muslim religion visually (by presence of the headscarf), which is a prominent (yet ordinary) cue, similar to an Arabic accent as a possible clue for ethnicity. We preselected 14 faces of young women that had relatively neutral appearance (i.e., non-indicative of their ethnicity), which were portrayed both with and without a headscarf. These women did not normally wear the headscarf and were approached individually by a researcher who was very familiar with headscarf use. Each woman was then photographed without and with a headscarf in front of the same white background. Different women were approached for the voice recordings in order to allow for the use of matched guise technique (MGT; Lambert, Hodgson, Gardner, & Fillenbaum, 1960). Speakers were 10 bilingual German–Persian and German-Arabic women who were recorded saying the same sentences both with a standard German accent and an Arabic accent. We decided to use the same speaker for both accent versions in order to control for the other voice characteristics (e.g., pitch, tone of the voice).

Additionally, this made visual and auditory stimuli comparable as both have the same faces and voices only with or without headscarf or accent, respectively. Because German native speakers were not able to differentiate a Persian from an Arabic accent, both were used. We tested both faces and voices for perceived attractiveness, prototypicality (both German and Arabic), and femininity; additionally, for voices we added a question about perceived accent strength and fluency (cf. Ryan, Carranza, & Moffie, 1977). All ratings were given on a 7-point Likert scale (1 = *not at all* to 7 = *very much*). The final sample consisted of 8 faces and 8 voices (from 16 different women) that were used in the experiment.

Face Selection

The 8 women used in this study were chosen as their ratings of attractiveness were rather similar when rated with and without wearing a scarf, $M_{\text{scarf}} = 4.44$ ($SD = 1.15$) and $M_{\text{no-scarf}} = 5.08$ ($SD = 0.99$), $t(28) = -4.32$, $p < .001$. Even though significant, this difference is due only to a headscarf bias, as the faces were the same in both cases. A similar bias was observed for perceived prototypicality of faces; again, participants used the headscarf as a prompt for higher Arabic prototypicality ($M_{\text{Arabic}} = 4.89$, $SD_{\text{Arabic}} = 1.23$; $M_{\text{German}} = 2.48$, $SD_{\text{German}} = .98$), $t(28) = -7.52$, $p < .001$; and consequently no headscarf for German prototypicality ($M_{\text{German}} = 4.21$, $SD_{\text{German}} = 1.20$; $M_{\text{Arabic}} = 2.70$, $SD_{\text{Arabic}} = 1.09$), $t(28) = -4.46$, $p < .001$. Interestingly, though faces were selected to be relatively neutral, they appeared more prototypically Arabic in the presence of the headscarf. Because in both cases participants rated the same person the this effect was not central. Similarly, the perceived femininity of faces reflected the same headscarf bias, $M_{\text{scarf}} = 5.42$ ($SD = 1.10$) and $M_{\text{no-scarf}} = 6.11$ ($SD = .80$), $t(28) = -4.03$, $p < .001$.

Voice Selection

A total of 8 female speakers were selected for the experiment. With regard to ratings of attractiveness, a similar bias as with faces was observed: Voices with a standard German

accent ($M_s = 5.42$, $SD_s = .66$) were perceived as more attractive than those speaking with an “Arabic” accent ($M_s = 4.57$, $SD_s = .83$), $t(28) = 6.36$, $p < .001$, even though the same speakers were rated in both cases. Specifically, in the present study we used Arabic and Persian accents. Even though Persian does not belong to the group of Arab languages the term Arabic accent was used in the pre-test as it was the most comprehensible for our participants. More importantly, speakers who spoke with a standard German accent were rated highly prototypical as German, not Arabic ($M_{German} = 5.80$, $SD_{German} = 1.04$ and $M_{Arabic} = 1.42$, $SD_{Arabic} = .48$), $t(28) = 2.07$, $p < .001$; and as more typically Arabic than German, with an Arabic accent ($M_{Arabic} = 3.65$, $SD_{Arabic} = 1.63$, and $M_{German} = 1.65$, $SD_{German} = .59$), $t(28) = -6.64$, $p < .001$. The somewhat lower prototypicality of Arabic accent can be explained with little familiarity of our participants with what constitutes an Arabic accent. More importantly speakers with an Arabic accent have been perceived significantly different from prototypical German speakers which was the crucial distinction for the present experiment. Not surprisingly, accent was perceived as stronger for speakers with an Arabic ($M = 6.03$, $SD = .59$) than with a standard German accent ($M = 1.84$, $SD = .81$), $t(28) = -24.17$, $p < .001$. The opposite pattern was observed for perceived fluency, where standard German ($M = 6.16$, $SD = .64$) was perceived as more fluent than the Arabic accent ($M = 4.05$, $SD = .99$), $t(28) = -1.55$, $p < .001$. Even though the accent strength was stronger (weaker for fluency) for the Arabic speakers they were intelligible and as previously explained the content of statements was always in standard German.

Appendix B

Homogeneity Test

Because the use of multinomial modelling assumes that the data are homogeneous across the sample, we conducted a test of homogeneity before proceeding with hypothesis testing (Klauer, 2010; Smith & Batchelder, 2010). The test was conducted on the baseline model using the TreeBUGS package for R, which allows for hierarchical multinomial-processing-tree modeling (Heck, Arnold, & Arnold, 2018). The parameter estimates together with standard deviations and confidence intervals suggest very little variation between participants (see Table B1). Indeed, when observing parameter estimates for each participant the same can be found (Parameter estimates for individual participants are available with the original data from this study at DOI: 1.17635/lancaster/researchdata/103). Therefore, we can conclude that the data obtained from our sample was homogeneous and consequently further analyses were conducted on overall frequencies.

Table B1

Parameter Estimates for the Homogeneity Test Including Confidence Intervals

Parameter	Mean	SD	2.50%	50%	97.50%
b	.20	.03	.14	.20	.27
cnG	.25	.04	.18	.26	.33
cnA	.04	.03	.00	.04	.12
chG	.11	.04	.03	.11	.18
chA	.16	.04	.08	.16	.24
DnG	.66	.04	.58	.66	.73
dnGf	.12	.07	.01	.11	.27
DnA	.63	.03	.57	.63	.68
dnA	.24	.07	.06	.24	.37

Parameter	Mean	SD	2.50%	50%	97.50%
dhG	.12	.07	.01	.12	.28
dhA	.34	.08	.17	.34	.50
enG	.32	.21	.02	.30	.75
enA	.84	.14	.41	.88	1.00
ehG	.67	.23	.13	.69	.98
ehA	.19	.11	.01	.17	.43
f	.51	.03	.46	.51	.58
hnG	.17	.09	.02	.17	.35
hnA	.24	.11	.03	.24	.46
hhG	.11	.08	.01	.09	.27
hhA	.43	.12	.19	.43	.66
x	.53	.03	.47	.53	.58

Note: the subscripts of different parameters indicate type of targets based on indicates the two categories (religion and ethnicity): n = no headscarf, h = headscarf, A = German with an Arabic accent, G = standard German.

References

- Heck, D. W., Arnold, N. R., & Arnold, D. (2018). Treebugs: An r package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*(1), 264-284. doi:10.3758/s13428-017-0869-7
- Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*(1), 70-98. <https://doi.org/10.1007/s11336-009-9141-0>
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, *60*(1), 44-51. <https://doi.org/10.1037/h0044430>
- Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of spanish-english bilinguals. *Language and Speech*, *20*(3), 267-273. <https://doi.org/10.1177/002383097702000308>
- Smith, J. B., & Batchelder, W. H. (2010). Beta-mpt: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology*, *54*(1), 167-183. <https://doi.org/10.1016/j.jmp.2009.06.007>