

Foreign-Looking Native-Accented People: More Competent When First Seen Rather Than
Heard?

Karolina Hansen*

University of Warsaw

Tamara Rakić

Lancaster University

Melanie C. Steffens

University of Koblenz-Landau

Version: August 2017

In press in Social Psychological and Personality Science

*Corresponding author.

University of Warsaw, Faculty of Psychology, Stawki 5/7, 00-183 Warszawa, Poland, +48 22

554 98 48, karolina.hansen@psych.uw.edu.pl

Abstract

Psychological research has neglected people whose accent does not match their appearance. Most research on person perception has focused on appearance, overlooking accents that are equally important social cues. If accents were studied, it was often done in isolation (i.e., detached from appearance). We examine how varying accent and appearance information about people affects evaluations. We show that evaluations of expectancy-violating people shift in the direction of the added information. When a job candidate looked foreign, but later spoke with a native accent, his evaluations rose and he was evaluated best of all candidates (Experiment 1a). However, the sequence in which information was presented mattered: when heard first and then seen, his evaluations dropped (Experiment 1b). Findings demonstrate the importance of studying the combination and sequence of different types of information in impression formation. They also allow predicting reactions to ethnically mixed people, who are increasingly present in modern societies.

Keywords: non-native speakers, face, voice, expectancy violations, stereotypes

Any person preparing for a job interview may have wondered how to make the best impression. If several pieces of information about a person are incongruent, the sequence of presenting them could play a role. Take, for example, a middle-eastern looking man who speaks with a standard accent. Is it better for him to stress ingroup language competence and only later reveal his foreign origin? Or should he start with his foreign appearance and then reveal his standard accent? Such cross-modal effects, though frequent in real life, are relatively little studied in psychology (see also Freeman & Ambady, 2011; Zuckerman, Miyake, & Hodgins, 1991). The present article aims at examining the influence of people's appearance and accents on observers' evaluations. In two experiments, we evoke expectations with either auditory or visual cues and then add the other type of information. This mirrors everyday life situations where only appearance or accent is available initially, for example, when seeing a silent person or when talking with someone on the phone.

Most research on person perception has focused on appearance, overlooking accents. Accents are at least equally important social cues. Two important theories, communication accommodation theory (Shepard, Giles, & Le Poire, 2001) and ethnolinguistic identity theory (ELIT, Giles & Johnson, 1981, 1987), show that language is a marker of identity. Language and accent are strong ingroup/outgroup markers and others' impressions are often based on it. Although ELIT does not address the role of appearance, it seems reasonable to assume that the role of accent in forming impressions of others could be especially pronounced when encountering people who speak in an unexpected way given their appearance. In such cases, accent can be an especially relevant cue for categorization and evaluation (Hansen, Rakić, & Steffens, 2017; Rakić, Steffens, & Mummendey, 2011).

To the best of our knowledge, only a few studies combined appearance and accent in forming impressions. Older studies used verbal stimuli of Blacks and Whites and showed

stronger effects of speech style than of race (Jussim, Coleman, & Lerch, 1987; McKirnan, Smith, & Hamayan, 1983). Recent research pitting accents against appearance demonstrated that accents play a bigger role for social categorization (Pietraszewski & Schwartz, 2014b; Rakić et al., 2011) and social evaluation (Hansen, Rakić, et al., 2017; Kinzler, Shutts, Dejesus, & Spelke, 2009; Rödin & Özcan, 2011). Although researchers have argued that appearance could sometimes be more important for social categorization than accent (Pietraszewski & Schwartz, 2014a), in all studies, accents were stronger social cues than appearance.

When people encounter others whose appearance and accent do not match, evaluations could be guided by the fact that these others violate one's expectations. As expectancy violation theory postulates, such violations should produce more extreme outcomes than situations matching expectations (Burgoon, 2009; Roese & Sherman, 2007). For example, if one had expected a conversation to be unpleasant, but it turned out to be pleasant, one would perceive it as even more pleasant than if one had already expected it to be pleasant (Burgoon & LePoire, 1993).

Such expectations can be evoked by ethnicity-related stereotypes. For example, in an American study Whites who spoke nonstandard English were viewed more negatively than Blacks who did, representing negative expectancy violations (Jussim et al., 1987). Conversely, another American study showed that Blacks with strong academic qualifications were evaluated as more competent than Whites with similar credentials, representing positive expectancy violations (Jackson, Sullivan, & Hodge, 1993). In a recent German study, Turkish-looking job candidates who spoke with a standard German accent were not only evaluated positively because of their standard accent, but they were evaluated even more positively than German-looking German-accented candidates (Hansen, Rakić, et al., 2017). Standard accents of foreign-looking candidates positively violated participants' negative expectations.

Although the above studies suggest that the evaluations are due to expectancy violations, these and other studies only assumed expectations; measuring them would be methodologically cleaner. The present study's first major contribution is that it tests expectations by showing one piece of information first, measuring evaluations, and then adding another piece. Our second contribution is to test how the sequence of presenting verbal and auditory information about a person influences evaluations.

We focused on competence evaluations and also explored hirability (for warmth and suggested salary, see Supplementary Material). Competence and warmth (agency and communion) are two main dimensions of person perception (Abele & Wojciszke, 2007; Fiske, Cuddy, Glick, & Xu, 2002). When people judge themselves or interdependent others, competence is seen as more important than warmth (Wojciszke & Abele, 2008). It has been consistently shown that foreign-accented speakers are perceived as less competent than native speakers, but results for warmth are mixed (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012). We studied these evaluations in Germany with Turks and Germans as alleged job candidates.

Current Research and Hypotheses

The current research tests expectancy violations theory with regard to accents. In Experiment 1a, we presented photographs of native- or foreign-looking job candidates and later added their voices that had accents violating (or not) the appearance-based expectations. In Experiment 1b the sequence of information was reversed.

As results for warmth are mixed (Fuertes et al., 2012) and there is little research on the perception of warmth based on Turkish accents in Germany, we did not formulate specific hypotheses except that expectancy violations also would happen. As hirability depends on competence and warmth evaluations (e.g., Rudman & Glick, 2001), and we could not formulate

hypotheses for warmth, we could not formulate strict hypotheses for hirability. However, we expected that hirability would replicate more closely competence rather than warmth findings.

In Experiment 1a where appearance was available first, we hypothesized that the addition of a Turkish accent would be a negative violation of expectations and should decrease competence ratings (Hypothesis 1a). Adding a German accent would evoke a positive violation and increase them (Hypothesis 1b). We expected these changes to be larger for incongruent than congruent job candidates (Hypothesis 1c). We expected that in the final evaluations among German-accented candidates those looking Turkish would be evaluated as even more competent than the German-looking ones (Hypothesis 1d). Corroborating negative expectancy violations, Turkish-accented German-looking candidates should be evaluated as less competent than Turkish-Turkish candidates (Hypothesis 1e).

In Experiment 1b where accents were available first, we expected that the addition of Turkish appearance should decrease competence ratings (Hypothesis 2a) and of German appearance – increase them (Hypothesis 2b). Again, for incongruent candidates such changes should be larger than for congruent candidates (Hypothesis 2c). Further, German-accented but (later revealed as) Turkish-looking candidates should be evaluated as less competent than German-German candidates (Hypothesis 2d), and Turkish-accented but German-looking candidates – as more competent than Turkish-Turkish candidates (Hypothesis 2e). We expected that the sequence of presenting candidates' appearance and accents would change evaluations, especially for incongruent candidates. We hypothesized that presenting a German-accented voice followed by a Turkish-looking face would cause a negative expectancy violation, so that the job candidate whose competence evaluations increased in Experiment 1a, would be evaluated as less competent (Hypothesis 3a). Conversely, adding a German-looking face to a Turkish-accented voice would evoke positive expectancy violations and increase competence ratings instead of

decreasing ratings (Hypothesis 3b). Furthermore, we explored hirability as well as warmth and suggested salary. We also explored whether hirability was mediated by candidates' competence and warmth and whether these potential indirect effects differed when appearance versus accent was presented first.

Method

Based on previous research on the power of accents as social cues (e.g., Hansen, Rakić, et al., 2017) we expected to obtain medium or large effect sizes. For Experiment 1a, we aimed at recruiting 50 or more participants (for the within-subject design) and reserved two lab days. We did not look at the data in-between. We decided a priori to exclude from analyses data from non-native speakers of German (Experiment 1a: 3, Experiment 1b: 1) and of participants who correctly guessed the experimental manipulation (Experiment 1b: 2).

Participants

Participants were undergraduate students from a German university. The final sample consisted of 60 participants (19 men, $M_{\text{age}} = 23.32$, $SD = 4.50$) in Experiment 1a and 54 in Experiment 1b (34 women, $M_{\text{age}} = 22.69$, $SD = 3.67$). Participants in Experiment 1b came from the same population as those in Experiment 1a and reported they had not taken part in a similar study. Samples did not differ demographically. Participants were compensated with either €2 and a chocolate bar or partial course credit. Given the within-subject design, the statistical power to detect medium effects ($f = .25$, Cohen, 1977) with $\alpha = .05$ and an assumed correlation of $r = .50$ between repeated measures was $1 - \beta = .97$ (Faul, Erdfelder, Lang, & Buchner, 2007).

Experimental Design

Both experiments had a 2 (appearance: German vs. Turkish) \times 2 (accent: standard German vs. German with a Turkish accent) \times 2 (time point: Time 1, only appearance/only accent vs. Time 2, appearance and accent) within-subject design. The experiments consisted of two evaluation blocks with eight job candidates in each. The first block included evaluations of four German-looking and four Turkish-looking faces (1a) or four German-accented and four Turkish-accented voices (1b). The second block included evaluations of two candidates out of each of the following four types: German appearance /German accent (GG, congruent), Turkish appearance /Turkish accent (TT, congruent), Turkish appearance/German accent (TG, incongruent), and German appearance/Turkish accent (GT, incongruent). Stimulus composition was counterbalanced and randomized: Any given voice (e.g., speaking standard German) was matched with a randomly selected congruent picture (German-looking person) in one version and with a randomly selected incongruent picture (Turkish-looking person) in the other. Candidates were presented in a random order.

We chose Turks as targets because they are the largest immigrant group in Germany (Federal Ministry of the Interior, 2007), stereotypically perceived as low in competence (Asbrock, 2010; Eckes, 2002). We used only male targets because differences in perceiving men and women are well-documented (e.g., Harper & Schoeman, 2003; O'Connell & Rotter, 1979) and the prototype of a Turk in Germany is a young man (e.g., Klingst & Drieschner, 2005).

Procedure, Materials, and Measures

After being welcomed, participants signed an informed consent form and were seated in front of a computer screen. First participants were asked to imagine they were helping in a recruitment process for a middle level manager position and they received résumés of many job candidates

(1a) or that candidates were calling them on the phone (1b). For each job candidate, participants were instructed to look at the photograph (1a) or listen to the voice (1b) and form an impression. All visual and auditory stimuli were selected after extensive pre-testing with regard to comparable attractiveness and pleasantness, contrasting German- and Turkish-typicality, and for voices additionally contrasting accent strength (see Supplementary Material). As the main dependent measure, we used short versions of the competence and warmth scales (e.g., Asbrock, 2010; Fiske et al., 2002) with the three items for competence (competent, competitive, and independent, all $\alpha \geq .84$) and warmth (likeable, warm, good-natured, $\alpha \geq .91$), all rated on a 7-point scale, 1 – *not at all* to 7 – *very much*. After this, participants indicated whether they would recommend to hire this candidate (1 – *definitely not* to 7 – *definitely yes*) and what salary they would suggest for him (answers needed to fit in-between 2000€ and 4000€ a month).¹ Then, participants were asked to imagine that the candidates came to the interview. Participants were instructed to evaluate the candidates again, but this time, half a second after seeing an already familiar face or hearing an already familiar voice, the second piece of information was added. All candidates said the same standard sentence in German (“Good morning, nice to meet you.”). At the end, participants answered a few demographic questions, provided their e-mail address for debriefing, were given their reward, thanked, and dismissed.²

¹ In the interest of brevity and clarity, we report in the main text competence and hirability results as we had hypotheses for those, and in the Supplementary Material, warmth and suggested salary. Although warmth analyses are reported in Supplementary Material, we refer to key findings in the main text where necessary.

² We also asked where, presumably, the candidates and their parents were born, and added open-ended questions about impressions of candidates (see Hansen, 2013). Contact with Turkish-origin people and motivation to respond without prejudice were assessed at the end and did not moderate findings.

Results

Competence

Appearance first (Experiment 1a). In order to test our hypotheses about shifts in competence ratings when accents are added to appearance, we conducted a 2 (appearance: German vs. Turkish) \times 2 (accent: German vs. Turkish) \times 2 (time point: Time 1, appearance vs. Time 2, appearance and accent) repeated measures ANOVA. These were followed by simple main effects tests of the a-priori hypotheses (see Supplementary Material for other effects).

Changes over time. The perceived competence of Turkish-looking candidates increased when they spoke standard German, $F(1,59) = 25.70$, $p < .001$, $\eta_p^2 = .30$, 95% CI [0.87, 0.38], in line with H1a and positively violated expectations (Figure 1). The evaluation of congruent German candidates also increased, $F(1,59) = 5.74$, $p = .02$, $\eta_p^2 = .09$, 95% CI [0.48, 0.04]. The competence of German-looking candidates decreased when they spoke with a Turkish accent, $F(1,59) = 13.75$, $p < .001$, $\eta_p^2 = .19$, 95% CI [-0.21, -0.70], confirming H1b and negatively violated expectations. The change in evaluations of congruent Turkish candidates was not significant, $F(1,59) = 2.69$, $p = .11$, $\eta_p^2 = .04$, 95% CI [0.06, -0.61]. The mean differences and effect sizes of change in evaluations for incongruent candidates were larger (Turkish appearance/German accent, $\Delta M = 0.60$, $\eta_p^2 = .30$ and German appearance/Turkish accent, $\Delta M = -0.45$, $\eta_p^2 = .19$) than for congruent candidates (German, $\Delta M = 0.27$, $\eta_p^2 = .09$ and Turkish, $\Delta M = -0.28$, $\eta_p^2 = .04$), supporting H1c.

Final evaluations. Comparing the final evaluations at Time 2 and corroborating the hypothesis about positively-violated expectations (H1d), Turkish-looking German-accented candidates were evaluated better than German-German candidates, $F(1,59) = 5.76$, $p = .02$, $\eta_p^2 = .09$, 95% CI [0.05, 0.53]. However, German-looking Turkish-accented candidates were not evaluated as less competent than Turkish-Turkish candidates, $F(1,59) = 1.27$, $p = .27$, $\eta_p^2 = .02$,

95% CI [-0.12, 0.44] (H1e). Thus, positive expectancy violations were visible both in changes of evaluations over time and in the final evaluations, but negative violations were visible only in changes over time.

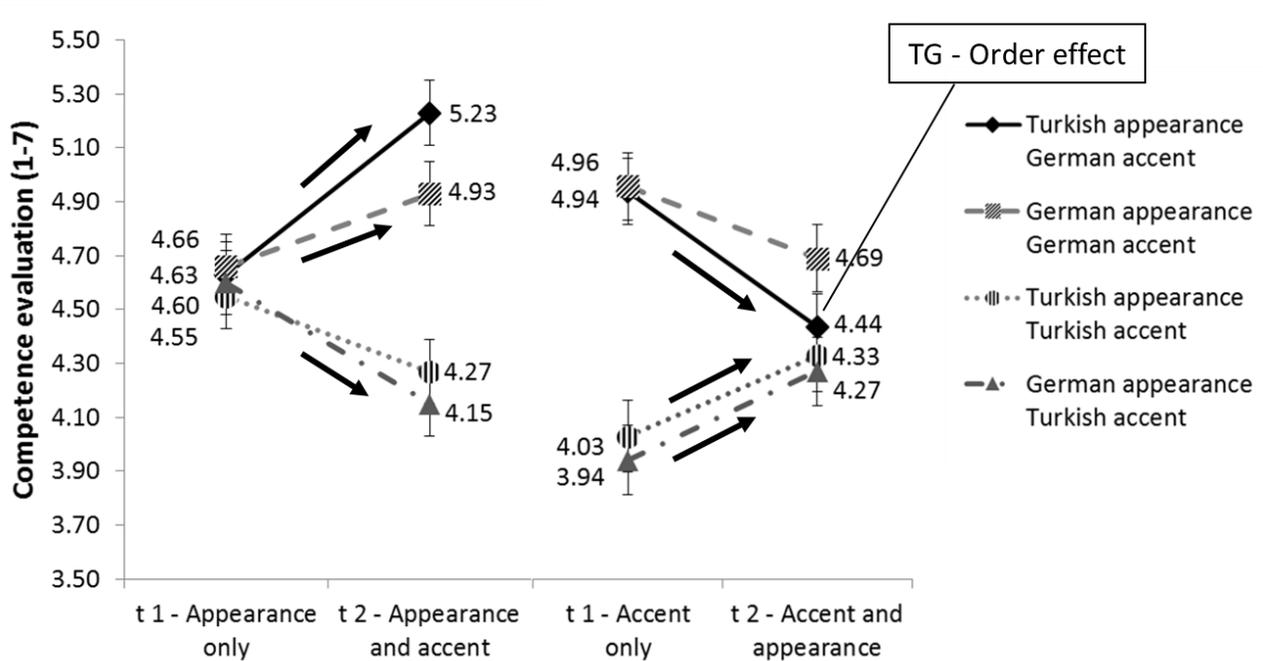


Figure 1. Shifts in mean competence evaluations by job candidate type in Experiment 1a (left) where appearance was presented first (t1) and accents were added (t2), and in Experiment 1b (right) where accents were presented first (t1) and appearance was added (t2). Note especially the difference in t2 evaluations of Turkish-looking but German-accented candidates. Error bars represent standard errors of the mean. Arrows indicate significant shifts of evaluations.

Accents first (Experiment 1b). An analogous analysis was carried out with accents first.

Changes over time. The perceived competence of German-accented candidates decreased when they looked Turkish, $F(1,53) = 14.38$, $p < .001$, $\eta_p^2 = .21$, 95% CI [-0.77, -0.24], confirming negatively violated expectations (H2a, Figure 1). The change in evaluations of congruent German candidates was not significant, $F(1,53) = 3.63$, $p = .06$, $\eta_p^2 = .06$, 95% CI [-0.55, 0.01]. The competence of Turkish-accented candidates increased when they looked

German, $F(1,53) = 5.78$, $p = .02$, $\eta_p^2 = .10$, 95% CI [0.06, 0.61], confirming positively violated expectations (H2b). However, the evaluation of congruent Turkish candidates also increased, $F(1,53) = 7.33$, $p = .009$, $\eta_p^2 = .12$, 95% CI [0.08, 0.52]. The effect sizes of change in evaluations for incongruent candidates were only partly larger ($\eta_p^2 = .21$ and $\eta_p^2 = .10$) than of congruent candidates ($\eta_p^2 = .06$ and $\eta_p^2 = .12$; H2c). The largest shift was noted for the Turkish-looking German-accented candidate ($\eta_p^2 = .21$), replicating the largest shift for this candidate in Experiment 1a.

Final evaluations. Turkish-looking German-accented candidates were not evaluated differently than Turkish-Turkish candidates, $F(1,53) = 0.67$, $p = .42$, $\eta_p^2 = .01$, 95% CI [-0.16, 0.38], not confirming H2d. German-looking Turkish-accented candidates were evaluated as less competent than German-German candidates confirming H2e, $F(1,53) = 5.58$, $p = .02$, $\eta_p^2 = .10$, 95% CI [0.06, 0.76].

Sequence effects. Comparing the data of both experiments (at Time 2), we found that Turkish-looking candidates who spoke standard German were evaluated significantly more competent when they were seen first than when they were heard first, $F(1,112) = 26.26$, $p < .001$, $\eta_p^2 = .19$, 95% CI [0.47, 1.06] (for GT, $F < 1$; TT, $F < 1$, GG, $F = 1.62$, $p = .21$, $\eta_p^2 = .01$, 95% CI [-0.13, 0.58]), corroborating H3a, but not H3b.

Hirability

Appearance first. Changes over time. Similarly as for competence, hirability evaluations of Turkish-looking candidates increased when they spoke standard German, $F(1,59) = 7.64$, $p = .008$, $\eta_p^2 = .12$, 95% CI [0.11, 0.70] (Figure 2). Evaluations of German-looking candidates also slightly increased when they spoke standard German, $F(1,59) = 4.01$, $p = .05$, $\eta_p^2 = .06$, 95% CI [0.00, 0.58]. Hirability of German-looking candidates, similarly as for competence, slightly decreased when they spoke with a Turkish accent, $F(1,59) = 3.37$, $p = .07$, $\eta_p^2 = .05$, 95% CI [-

0.59, 0.03]. At the same time the evaluations of Turkish-looking candidates, similarly as for warmth (see Supplementary Material), slightly *increased* when they spoke with a Turkish accent, $F(1,59) = 3.59, p = .06, \eta_p^2 = .06, 95\% \text{ CI} [-0.02, 0.58]$. These hirability results appear to reflect the combination of perceived competence and warmth of the candidates.

Final evaluations. German-looking Turkish-accented candidates (incompetent and moderately warm) were perceived as less hirable than all other candidates, $F_s \geq 10.41, p_s \leq .002, \eta_p^2 \geq .15$. There were no differences between the other candidates, $F_s < 1$.

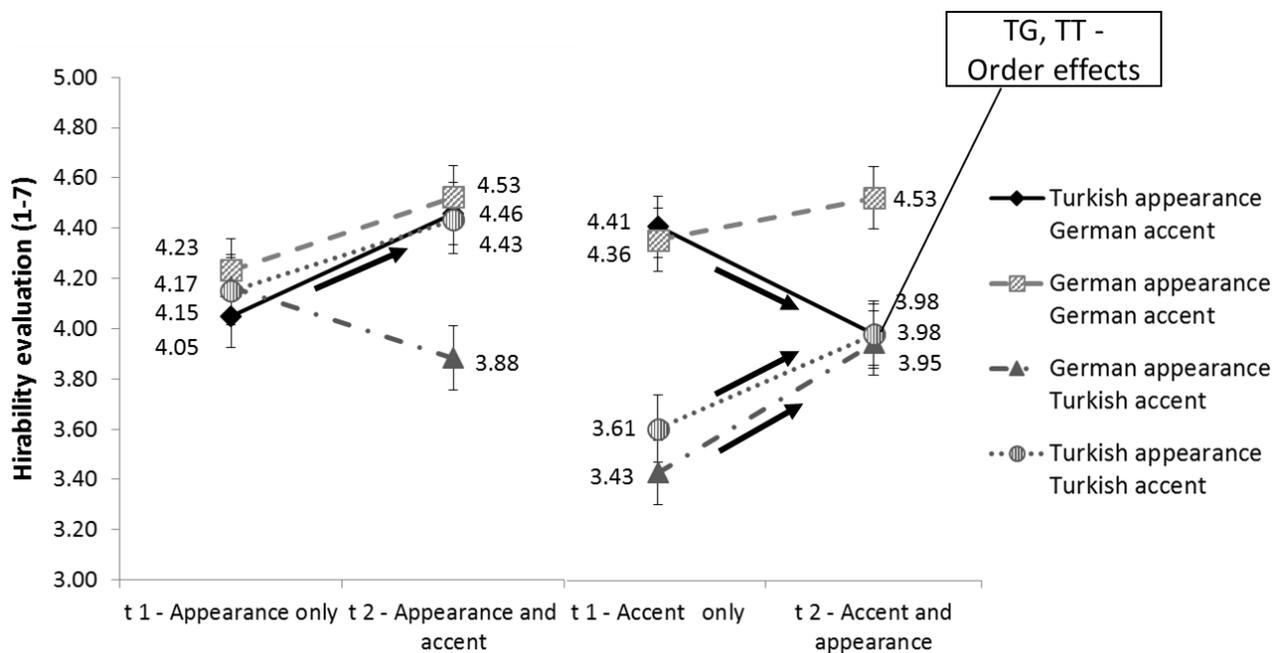


Figure 2. Shifts in mean hirability evaluations by job candidate type in Experiment 1a (left) where appearance was presented first (t1) and accents were added (t2), and in Experiment 1b (right) where accents were presented first (t1) and appearance was added (t2). Note especially the difference in t2 evaluations of Turkish-looking but German-accented candidates, as well as Turkish-Turkish candidates. Error bars represent standard errors of the mean. Arrows indicate significant shifts of evaluations.

Accents first. Changes over time. Similarly as for competence, the perceived hirability of German-accented candidates decreased when they looked Turkish, $F(1,53) = 10.89, p = .002, \eta_p^2 = .16, 95\% \text{ CI} [-0.83, -0.20]$. Comparably, the change for congruent German candidates was not significant, $F(1,53) = 1.17, p = .28, \eta_p^2 = .02, 95\% \text{ CI} [-0.48, 0.14]$. As for competence, hirability of Turkish-accented candidates increased when they looked German, $F(1,53) = 6.36, p = .02, \eta_p^2 = .10, 95\% \text{ CI} [0.09, 0.77]$, and so did the evaluation of congruent Turkish candidates, $F(1,53) = 6.30, p = .02, \eta_p^2 = .10, 95\% \text{ CI} [0.08, 0.68]$.

Final evaluations. German-looking German-accented candidates were perceived as more hireable than all other candidates, $F_s \geq 9.65, p_s \leq .003, \eta_p^2_s \geq .15$ (other $F_s < 1$).

Sequence effects. The two sequences of presentation caused, as for competence, different evaluations of Turkish-looking German-accented candidates: they were evaluated as more hireable when seen first than when heard first, $F(1,112) = 6.92, p = .01, \eta_p^2 = .06, 95\% \text{ CI} [0.12, 0.86]$. There was a similar effect for Turkish-Turkish candidates, $F(1,112) = 7.15, p = .009, \eta_p^2 = .06, 95\% \text{ CI} [0.14, 0.84]$, (GT, $F < 1$; GG, $F < 1$). The sequence effect for Turkish-looking German-accented candidates was also replicated for warmth (see Supplementary Material).

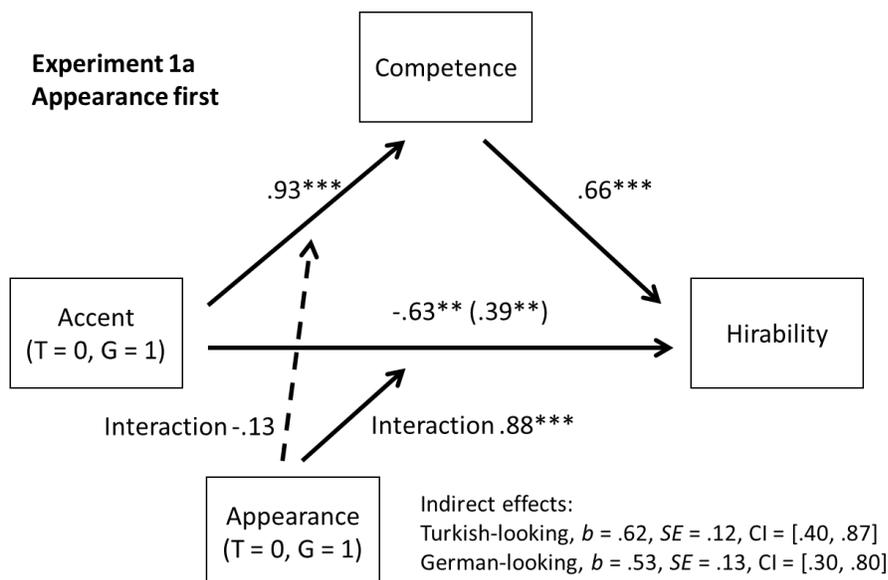
Differences in Indirect Effects of Competence on Hirability

As Table S3 in the Supplementary Material shows, correlations between variables were not too high to analyze indirect effects.

Appearance first. To explore whether sequence effects in competence transmit to hirability, we tested a moderated mediation model with accent as the independent variable, appearance as a moderator, hirability (at Time 2) as a dependent variable, and competence (Time 2) as a mediator (Model 8 in Hayes, 2013). We used 95% bias-corrected (BC) bootstrapped confidence intervals based on 5,000 bootstrap samples. To see whether the sequence effects of competence and

hirability surface in a joint analysis, we were especially interested in the comparison of indirect effects for different types of targets (for other effects, see Supplementary Material).

The analysis showed that disregarding competence, German-accented job candidates were perceived as more hireable, $b = .39$, $SE(\text{boot}) = .15$, $CI = [.10, .68]$ (Figure 3, upper panel). However, when including competence, Turkish-accented job candidates were perceived as more hireable, $b = -.63$, $SE(\text{boot}) = .19$, $CI = [-.99, -.26]$. The observed suppression effect suggests that there are two contrary mechanisms, possibly higher competence of German-accented, but higher warmth of Turkish-accented candidates (see "negative suppression" in Conger, 1974; MacKinnon & Fairchild, 2009). Indirect effects of accent on hirability through competence were significant for both types of candidates' appearance: Turkish-looking candidates when speaking standard German were perceived as more competent and thus more hireable, $b = .62$, $SE(\text{boot}) = .12$, $CI = [.40, .87]$, and German-looking candidates when speaking standard German were also perceived as more competent and thus more hireable, $b = .53$, $SE(\text{boot}) = .13$, $CI = [.30, .80]$.



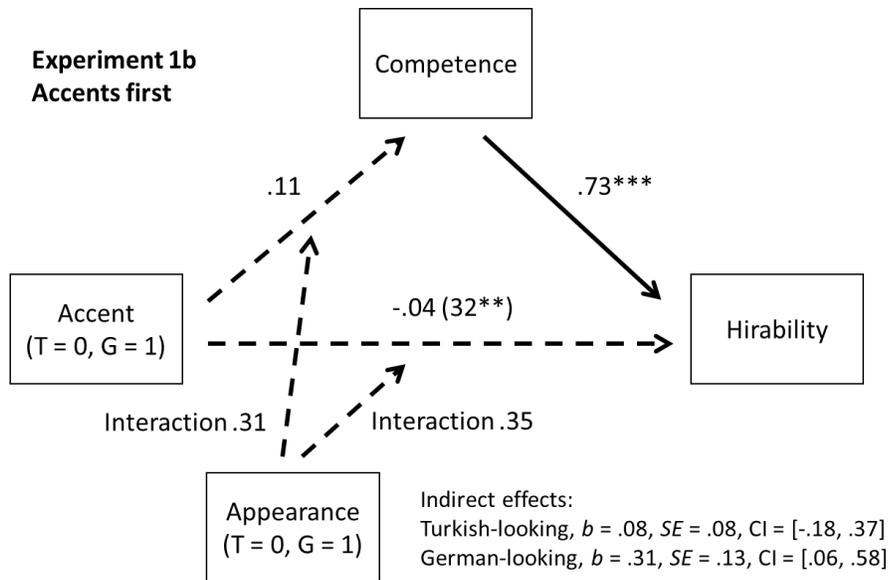


Figure 3. Indirect effects of accent and appearance on hirability via competence in Experiment 1a where accents were added to appearance (upper panel) and 1b where appearance was added to accents (lower panel).

Accent first. When accents were available from the beginning and appearance was added, disregarding competence, German-accented job candidates were perceived as more hireable, $b = .32$, $SE(\text{boot}) = .14$, $CI = [.04, .61]$. When including competence, the effect disappeared, $b = -.04$, $SE(\text{boot}) = .15$, $CI = [-.34, .25]$ (Figure 3, lower panel). The indirect effect of accent on hirability through competence was significant only for German appearance: German-accented candidates when also German-looking were perceived as more competent and thus more hireable, $b = .31$, $SE(\text{boot}) = .13$, $CI = [.06, .58]$, but for German-accented Turkish-looking candidates competence did not explain hirability, $b = .08$, $SE(\text{boot}) = .08$, $CI = [-.18, .37]$. In sum, the sequence of presenting candidates' faces and voices influenced not only the final evaluations, but also the process of competence ascription and the following recommendations to hire this person. The sequence mattered especially, again, for Turkish-looking targets: When first seen and then heard speaking standard German (vs. with a Turkish accent), they were evaluated as more (less)

competent and more (less) hireable, but when first heard (with any accent) and then seen as Turkish-looking, this effect was not observed. For German-looking targets the sequence did not matter: They were always perceived as more competent when speaking standard German (vs. with a Turkish accent) and this always lead to their higher hirability.

General Discussion

By considering both visual (appearance) and auditory (accent) cues, the present research contributes to understanding the influence of these cues on impression information. The sequence of presenting incongruent cues can also matter and the same person when seen first, can be evaluated differently than when heard first. When Turkish-looking German-accented candidates were first seen and heard later (Experiment 1a), their accent positively violated appearance-based expectations and they were evaluated as the most competent of all candidates. This replicates previous research showing similar effects of positive expectancy violations (Hansen, Rakić, et al., 2017; Hansen, Steffens, Rakić, & Wiese, 2017). However, when the standard accent was presented first and was followed by Turkish appearance, a negative expectancy violation and worse competence (and hirability) evaluation occurred (Experiment 1b).

Indirect effects analyses showed that when appearance was presented first, the mechanism of recommending candidates for a job was similar for Turkish- and German-looking candidates: when speaking standard German they were perceived as more competent and thus more hireable. However, when they were heard first, this mechanism worked only for German-, but not for Turkish-looking candidates. It seems that for congruent people it does not matter whether they are first seen or heard, but for incongruent people it does. Practically speaking, foreign-looking standard speakers should first reveal their provenance and later their speech.

Future research should test the boundary conditions of these sequence effects. For instance, would it suffice to present a foreign-sounding name first? Also, future research could

manipulate background information about the minority group. It could be studied whether, for example, presenting an anti-Turkish text could lower the already negative expectations and lead to an even greater positive surprise when a Turkish-looking person would speak standard. As for any ethnicity-related research, results may depend on the cultural context where the study is conducted. However, even if the specific accents or appearance in different cultures were associated with different stereotypes, the sequence effect could be more universal.

Interestingly, besides the expected sequence effects for Turkish-looking German-accented candidates both on competence and hirability, congruent Turkish candidates were also perceived as more hireable (but not more competent) when first seen rather than heard. Their hirability evaluations were unexpectedly high in Experiment 1a. A possible alternative explanation of this (and possibly other) results could be the motivation to control prejudice. Especially in the case of visual information or in the case of hiring of stereotypical immigrants (Turkish-looking, Turkish-accented), people may have corrected their responses to Turkish-looking faces, but, in contrast, they perceived accents as a reasonable, not discriminatory cue for evaluations. Although there is evidence that accent discrimination is perceived as a legitimate evaluation (Hansen & Dovidio, 2016; Souza, Pereira, Camino, Lima, & Torres, 2016), motivation to control prejudice or its interaction with other factors did not modify the present findings ($Bs < 0.09$, $ps > .38$).

In Experiment 1a, as in previous research (Hansen, Steffens, et al., 2017), participants were surprised by incongruent people. However, the competence and hirability results of Experiment 1b showed also shifts for congruent Turkish people. This may suggest that general cue-related and sequence-related processes similar to a dilution effect (de Vries, Terwel, & Ellemers, 2014; Nisbett, Zukier, & Lemley, 1981) or regression to the mean (Nesselrode, Stigler, & Baltes, 1980) also influenced evaluations. Even if these effects obscured effects of expectancy violations in Experiment 1b, based on Experiment 1a and on previous EEG research,

we assume that incongruent job candidates evoked participants' surprise (Hansen, Steffens, et al., 2017).

We suggest treating and measuring expectancy violations as a dynamic process with relative differences between what was expected and how the impression changed in the presence of a new piece of information. Our approach allows obtaining stronger evidence for expectancy violation theory and detecting rises or drops in evaluations that can have potentially important social consequences.

The current research suggests that researchers should pay more attention to the interactions of appearance, accent, and other cues in impression formation. It also underscores the practical importance of changes in evaluations, which can be especially frequent when people encounter others whose different attributes do not match. Reactions to and evaluations of such people have been little studied. With our research, we hope to pave an avenue for future research on the perception of such incongruent people.

Acknowledgements

We thank Claudia Niedlich for help in data collection, Wiktor Soral for help with converting the dataset, as well as the editor and an anonymous reviewer for their comments on a previous version of this manuscript.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship or the publication of this article.

Funding

The current research was supported by the ProExzellenz program of the state of Thuringia, Polish National Science Centre (NCN Fuga DEC-2013/08/S/HS6/00573), and German Research Foundation (DFG STE 938/10-2; FOR 1097).

References

- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751-763.
doi:10.1037/0022-3514.93.5.751
- Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology, 41*(2), 76–81.
- Burgoon, J. K. (2009). Expectancy violations theory. In E. Griffin (Ed.), *A first look at communication theory* (pp. 84-97). New York: McGraw-Hill.
- Burgoon, J. K., & LePoire, B. A. (1993). Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. *Human Communication Research, 20*(1), 67-96.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised ed.). Hillsdale, NJ: Erlbaum.
- Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and psychological measurement, 34*(1), 35-46.
- de Vries, G., Terwel, B. W., & Ellemers, N. (2014). Spare the details, share the relevance: The dilution effect in communications about carbon dioxide capture and storage. *Journal of Environmental Psychology, 38*(0), 116-123. doi:10.1016/j.jenvp.2014.01.003
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles, 47*(3-4), 99-114. doi:10.1023/A:1021020920715
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Federal Ministry of the Interior. (2007). Deutsche und Ausländer mit Migrationshintergrund [Germans and foreigners with an immigrant background]. Retrieved from

<http://www.en.bmi.bund.de/>

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902.

doi:10.1037//0022-3514.82.6.878

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of personal construal.

Psychological Review, 118(2), 247-279. doi:10.1037/a0022327

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, 42(1), 120-133. doi:10.1002/ejsp.862

Giles, H., & Johnson, P. (1981). The role of language in ethnic group relations. In J. C. Turner & H. Giles (Eds.), *Intergroup behavior* (pp. 199–243). Oxford, England: Blackwell.

Giles, H., & Johnson, P. (1987). Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language*, 68(1), 69-99. doi:10.1515/ijsl.1987.68.69

Hansen, K. (2013). *Interactions of accent and appearance in social categorization, impression formation, and economic decisions*. (PhD Thesis), Friedrich Schiller University, Jena,

Germany. Retrieved from http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-27328/Diss/HansenKarolina_Dissertation.pdf

Hansen, K., & Dovidio, J. F. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural diversity & ethnic minority psychology*.

- Hansen, K., Rakić, T., & Steffens, M. C. (2017). Competent and warm? How mismatching appearance and accent influence first impressions. *Experimental Psychology*. doi:10.1027/1618-3169/a000348
- Hansen, K., Steffens, M. C., Rakić, T., & Wiese, H. (2017). When appearance does not match accent: Neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsw148
- Harper, M., & Schoeman, W. J. (2003). Influences of gender as a basic-level category in person perception on the gender belief system. *Sex Roles, 49*(9), 517-526.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology, 65*(1), 69-84. doi:10.1037/0022-3514.65.1.69
- Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 52*(3), 536-546. doi:10.1037/0022-3514.52.3.536
- Kinzler, K. D., Shutts, K., Dejesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623-634. doi:10.1521/soco.2009.27.4.623
- Klingst, M., & Drieschner, F. (2005). Raus aus dem Ghetto [Get out of the ghetto]. *Die Zeit, 46*. Retrieved from http://www.zeit.de/2005/46/dt_Integration
- MacKinnon, D., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science, 18*(1), 16-20.

- McKirnan, D. J., Smith, C. E., & Hamayan, E. V. (1983). A sociolinguistic approach to the belief-similarity model of racial attitudes. *Journal of Experimental Social Psychology*, *19*(5), 434-447. doi:10.1016/0022-1031(83)90021-5
- Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin*, *88*(3), 622-637.
- Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*(2), 248-277.
- O'Connell, A. N., & Rotter, N. G. (1979). The influence of stimulus age and sex on person perception. *Journal of Gerontology*, *34*(2), 220-228.
- Pietraszewski, D., & Schwartz, A. (2014a). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, *35*(1), 51-57. doi:10.1016/j.evolhumbehav.2013.09.005
- Pietraszewski, D., & Schwartz, A. (2014b). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, *35*(1), 43-50. doi:10.1016/j.evolhumbehav.2013.09.006
- Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16-29. doi:10.1037/a0021522
- Rödin, M., & Özcan, G. (2011). Is it how you look or speak that matters? An experimental study exploring the mechanisms of ethnic discrimination (Research Paper No. 2011:12). Retrieved from http://econpapers.repec.org/RePEc:hhs:sunrpe:2011_0012

- Roese, N. J., & Sherman, J. W. (2007). Expectancy. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2 ed., pp. 91-115). New York: Guilford Press.
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*(4), 743-762.
- Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. In W. P. Robinson & H. Giles (Eds.), *The New Handbook of Language and Social Psychology* (2nd ed., pp. 33-56). Sussex, England: John Wiley & Sons.
- Souza, L. E. C., Pereira, C. R., Camino, L., Lima, T. J. S., & Torres, A. R. R. (2016). The legitimizing role of accent on discrimination against immigrants. *European Journal of Social Psychology, 46*(5), 609-620.
- Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology, 38*(7), 1139-1147.
- Zuckerman, M., Miyake, K., & Hodgins, H. S. (1991). Cross-channel effects of vocal and physical attractiveness and their implications for interpersonal perception. *Journal of Personality and Social Psychology, 60*, 545-554. doi:10.1037/0022-3514.60.4.545

Author Biographies

Karolina Hansen is an assistant professor at the University of Warsaw. Her research interests are in the fields of social psychology, sociolinguistics, and intercultural psychology, and include topics such as language and accent attitudes, stereotyping, gender-fair language, linguistic biases, and intercultural differences in social cognition.

Tamara Rakić is a lecturer at Lancaster University, UK. In 2011/12, she was a *Visiting Scholar* at University of California, Santa Barbara, USA. Her research interests include the influence of accents on person perception and categorization, and relation between language and social identity.

Melanie C. Steffens holds the position of full professor for Social and Economic Psychology at the Faculty of Psychology at the University of Koblenz-Landau (since 2013). Her research interests comprise implicit and explicit attitudes and stereotypes of social groups, social categorization and impression formation, in addition to memory phenomena.

Supplementary Material

Pre-Tests and Selection of Stimulus Materials

To ensure that the stimuli are perceived as typical for their respective groups, all stimuli were pre-tested by asking (two separate questions) how typically German and how typically Turkish the targets appeared or sounded. Additionally, to avoid the “What is beautiful is good” phenomenon (Dion, Berscheid, & Walster, 1972; Zuckerman & Driver, 1989) all stimuli were pre-tested for attractiveness and pleasantness. Moreover, linguistically lay people can reliably rate varying degrees of accent strength, and the stronger the accent, the more negative the evaluations (Ryan, Carranza, & Moffie, 1977). Therefore, audio stimuli were also pre-tested for accent strength.

The pre-test sample consisted of 29 participants (13 men, $M_{\text{age}} = 22.73$, $SD = 3.42$). The pre-test was conducted as a separate, independent study with participants who did not participate in any of the main experiments but were from the same population.

In the pre-test participants sat individually in front of computer screens and were presented with two blocks of stimuli: faces and voices. After each face or voice, participants answered questions about its attractiveness, pleasantness, and typicality, on 7-point scales ranging from 1 – *not at all* to 7 – *very much*. Both the stimuli and the subsequent questions appeared in random order. Voices were also evaluated regarding the strength of their foreign accent (1 – *no foreign accent at all* to 7 – *very strong foreign accent*).

Faces

We used portrait photographs of faces available in two online scientific databases (Langner et al., 2010; Minear & Park, 2004) and we added several photographs of Turkish men that we took ourselves. In total, we selected 36 photographs of faces for pre-testing (21 Turkish-looking, 15

German-looking). The format of the pictures was standardized and they were all converted into black and white. All photographs were of young men with a neutral facial expression, dressed neutrally, without glasses, and who did not have any stereotypical characteristics (e.g., no long moustache or beard for Turkish targets).

From the pre-tested faces, we selected four German- and four Turkish-looking faces, which were typical for their respective groups. German-looking faces were much more typically German than Turkish, $t(28) = 14.65, p < .001$. Analogously, Turkish faces were more typically Turkish than German, $t(28) = -8.29, p < .001$. Selected faces were moderately attractive and pleasant (descriptive statistics can be found in Table S1). The German- and Turkish-looking faces were similarly attractive, $t < 1$, and similarly pleasant, $t < 1$.

Table S1

Descriptive Statistics of Faces and Voices Selected for Experiment 1a and Experiment 1b.

	Faces		Voices	
	$M(SD)_{\text{German}}$	$M(SD)_{\text{Turkish}}$	$M(SD)_{\text{German}}$	$M(SD)_{\text{Turkish}}$
Attractiveness	3.02 (1.14)	2.97 (1.05)	3.24 (1.32)	3.09 (1.00)
Pleasantness	4.47 (0.89)	4.14 (1.05)	4.20 (1.02)	4.21 (0.71)
Typical German	5.33 (1.28)	2.27 (0.73)	5.30 (1.28)	2.27 (0.70)
Typical Turkish	1.34 (0.53)	4.36 (1.24)	1.34 (0.53)	4.36 (1.24)
Accent strength			1.30 (0.52)	4.80 (1.16)

Voices

Short voice samples of young German and Turkish native speakers were recorded. The speakers said a neutral everyday phrase, “Good morning, nice to meet you.” (In German: “Guten Tag, es freut mich, dass wir uns kennen lernen.”). By having all speakers say the same sentence any influence of content of the statement was excluded and it also ensured that accented sentences

were not more difficult to understand. The speakers were briefly trained and several versions were recorded of each speaker. Speech rate was kept constant; all the speakers said the statements at a medium speed so that the voice samples were three seconds long. All speakers had a typical male timbre of voice.

Similarly to faces, from the pre-tested voices we selected four German and four Turkish voices, which were typical for their group. German voices were much more typically German than Turkish, $t(28) = 14.65, p < .001$. Analogously, Turkish voices were more typically Turkish than German, $t(28) = -8.29, p < .001$. German voices were perceived to speak with no accent and Turkish voices to speak with a moderately strong one (Table S1). The difference in accent strength between German and Turkish voices was significant, $t(28) = -17.21, p < .001$, ensuring an effective manipulation. The voices were moderately attractive and pleasant. The German and Turkish voices were similarly attractive, $t < 1$, and similarly pleasant, $t < 1$.

Additional Results

Table S2.

ANOVA Results for Competence, Hirability, Salary, and Warmth in Experiment 1a (with Appearance First) and 1b (with Accents First).

1a Appearance first	Competence			Hirability			Salary			Warmth		
	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
Appearance	1.11	.30	.02	0.38	.54	.01	0.13	.72	.00	0.92	.34	.02
Accent	39.33	.00	.40	2.53	.12	.04	0.69	.41	.01	0.26	.61	.00
Time point	0.39	.54	.01	5.12	.03	.08	1.56	.22	.03	0.00	.99	.00
Appearance * accent	0.06	.81	.00	4.73	.03	.07	8.21	.01	.12	0.80	.38	.01
Appearance * time	4.77	.03	.07	7.14	.01	.11	6.95	.01	.11	12.76	.00	.18
Accent * time	27.34	.00	.32	4.82	.03	.08	5.47	.02	.08	0.12	.73	.00
Appearance * accent * time	0.56	.46	.01	2.13	.15	.03	2.17	.15	.04	0.94	.34	.02

1b Accent first	Competence			Hirability			Salary			Warmth		
	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
Appearance	0.15	.70	.00	0.62	.43	.01	0.18	.67	.00	0.74	.39	.01
Accent	39.04	.00	.42	19.40	.00	.26	25.33	.00	.31	6.24	.02	.11
Time point	0.25	.62	.00	4.16	.05	.07	0.42	.52	.01	5.26	.03	.09
Appearance * accent	1.58	.21	.03	3.97	.05	.07	2.92	.09	.05	5.71	.02	.10
Appearance * time	1.01	.32	.02	5.74	.02	.09	4.05	.05	.07	0.74	.39	.01
Accent * time	30.97	.00	.37	10.40	.00	.16	10.62	.00	.16	11.98	.00	.18
Appearance * accent * time	0.69	.41	.01	2.59	.11	.04	10.22	.00	.15	3.25	.08	.06

Note: Experiment 1a: Hypothesis $df = 1$, error $df = 59$; 1b: Hypothesis $df = 1$, error $df = 53$.

Warmth

Appearance first (Experiment 1a). Changes over time. Similar analyses as for competence showed that the warmth evaluations of Turkish-looking candidates, as for warmth and hirability, increased when they spoke standard German, $F(1,59) = 7.28$, $p = .009$, $\eta_p^2 = .11$, 95% CI [0.09, 0.58] (Figure S1). However, German-looking candidates appeared slightly less warm when they spoke standard German, $F(1,59) = 4.02$, $p = .05$, $\eta_p^2 = .06$, 95% CI [0.001, 0.55]. Warmth of German-looking candidates did not change when they spoke with a Turkish accent, $F(1,59) = 2.78$, $p = .10$, $\eta_p^2 = .05$, 95% CI [-0.04, 0.45]. The change in evaluations of congruent Turkish

candidates also was non-significant, $F(1,59) = 0.86, p = .36, \eta_p^2 = .01, 95\% \text{ CI} [-0.48, 0.18]$.

Thus, only the German accent evoked shifts in warmth that depended on appearance: Turkish-looking candidates became warmer with a German accent, but German-looking candidates became slightly less warm.

Final evaluations. The only significant difference between candidates in the final evaluations was between these two German-accented candidates, $F(1,59) = 12.17, p = .001, \eta_p^2 = .17, 95\% \text{ CI} [0.19, 0.71]$. In general, the results for warmth showed shifts in evaluations that were weaker and different than for competence.

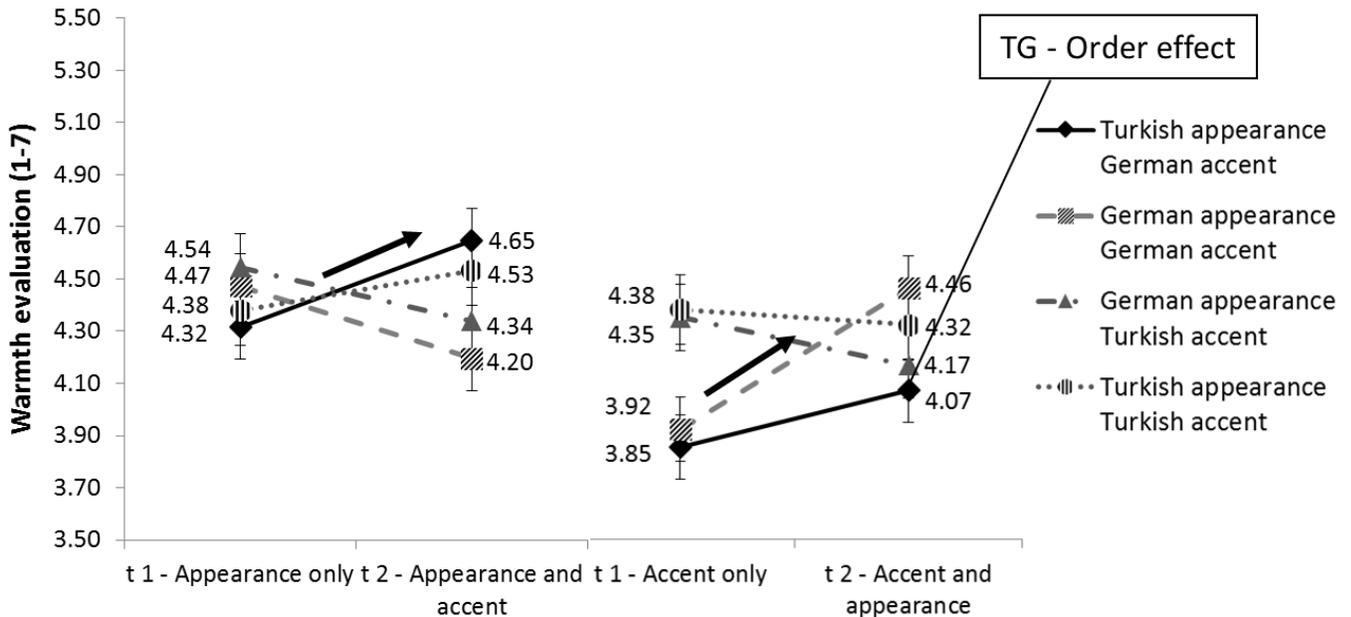


Figure S1. Shifts in mean warmth evaluations by job candidate type in Experiment 1a (left) where appearance was presented first (t1) and accents were added (t2), and in Experiment 1b (right) where accents were presented first (t1) and appearance was added (t2). Note especially the difference in t2 evaluations of Turkish-looking but German-accented candidates. Error bars represent standard errors of the mean. Arrows indicate significant shifts of evaluations.

Accents first (Experiment 1b). Changes over time. The perceived warmth of German-accented candidates increased slightly when they looked Turkish, $F(1,53) = 3.04, p = .09, \eta_p^2 = .05, 95\%$

CI [-0.03, 0.47]. Warmth of German-accented candidates increased more when they looked German, $F(1,53) = 12.09, p = .001, \eta_p^2 = .19, 95\% \text{ CI } [0.23, 0.85]$. Warmth of neither Turkish-accented candidate changed significantly (German-looking: $F(1,53) = 2.46, p = .12, \eta_p^2 = .04, 95\% \text{ CI } [-0.05, 0.42]$, Turkish-looking: $F(1,53) = 0.37, p = .55, \eta_p^2 = .01, 95\% \text{ CI } [-0.14, 0.25]$).

Final evaluations. Final evaluations of congruent German candidates were warmer than of German-accented Turkish-looking candidates, $F(1,53) = 8.96, p = .004, \eta_p^2 = .15, 95\% \text{ CI } [0.13, 0.66]$. Among Turkish-looking candidates Turkish-accented ones were perceived as warmer than German-accented ones, $F(1,53) = 5.72, p = .02, \eta_p^2 = .10, 95\% \text{ CI } [0.04, 0.47]$. Other comparisons were not significant. In sum, standard German was perceived as less warm than Turkish-accented German.

Sequence effects. The two sequences of presentation caused, again, different evaluations of Turkish-looking candidates who spoke standard German: They were evaluated warmer when seen first than when heard first, $F(1,112) = 9.44, p < .001, \eta_p^2 = .11, 95\% \text{ CI } [0.26, 0.89]$ (for GT, $F < 1$; TT, $F = 1.50, p = .22, \eta_p^2 = .01$; GG, $F = 2.48, p = .12, \eta_p^2 = .02, 95\% \text{ CI } [-0.60, 0.07]$).

Suggested salary

Appearance first. Changes over time. Similarly as for competence and hirability, suggested salary of Turkish-looking candidates increased when they spoke standard German, $F(1,59) = 5.56, p = .02, \eta_p^2 = .09, 95\% \text{ CI } [17.96, 218.76]$ (Figure S2). Salary of congruent German candidates did not change significantly, $F(1,59) = 2.53, p = .04, \eta_p^2 = .06, 95\% \text{ CI } [-23.41, 205.47]$. Salary of congruent Turkish candidates also did not change when they spoke with a Turkish accent, $F(1,59) = 1.78, p = .19, \eta_p^2 = .03, 95\% \text{ CI } [-35.53, 177.08]$. Hirability of German-looking candidates, similarly as for competence, decreased when they spoke with a Turkish accent, $F(1,59) = 4.40, p = .04, \eta_p^2 = .07, 95\% \text{ CI } [5.68, 239.621]$. These salary results replicate

competence and hirability results and again show that evaluations of incongruent targets shifted when in addition to a known appearance their voice could be heard.

Final evaluations. German-looking Turkish-accented candidates (least competent, moderately warm and least hireable) were suggested a lower salary than all other candidates, $F_s \geq 8.98$, $p_s \leq .004$, $\eta_p^2 \geq .13$. There were no differences between the other candidates, $F_s < 1$.

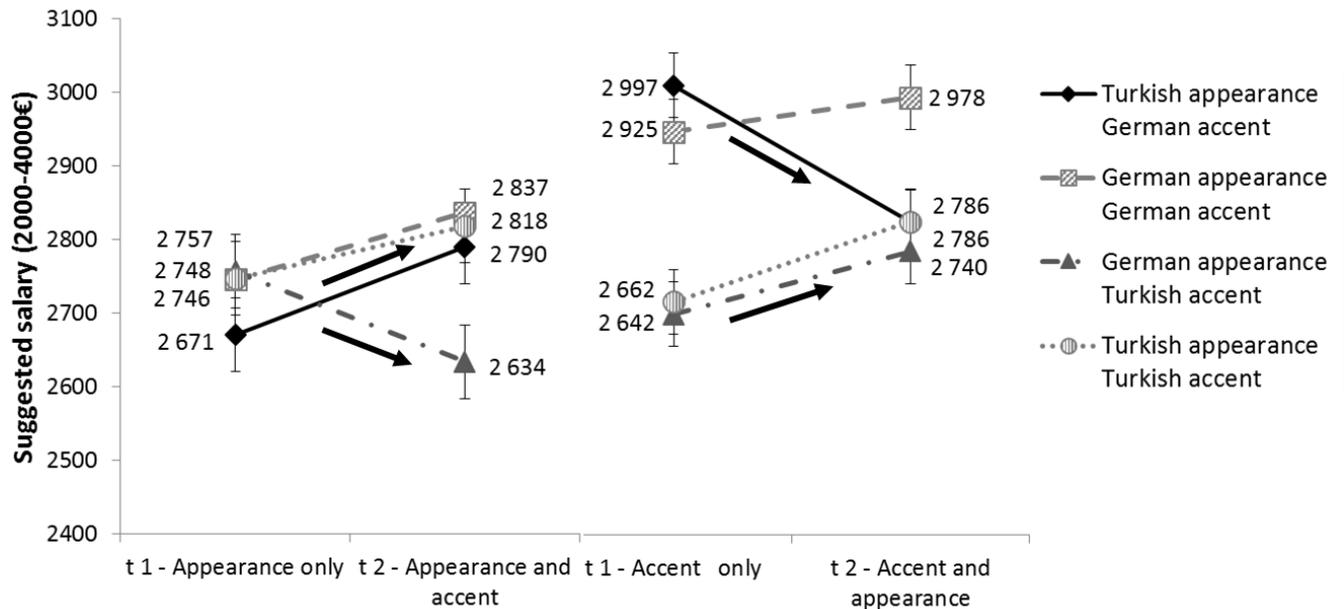


Figure S2. Shifts in mean suggested salary by job candidate type in Experiment 1a (left) where appearance was presented first (t1) and accents were added (t2), and in Experiment 1b (right) where accents were presented first (t1) and appearance was added (t2). Error bars represent standard errors of the mean. Arrows indicate significant shifts of evaluations.

Accents first. Changes over time. Similarly as for competence and hirability, salary of German-accented candidates decreased when they looked Turkish, $F(1,53) = 11.03$, $p = .002$, $\eta_p^2 = .17$, 95% CI [-338.69, -83.82]. The change for congruent German candidates was not significant, $F(1,53) = 1.02$, $p = .32$, $\eta_p^2 = .02$, 95% CI [-158.84, 52.52]. As for competence and hirability, salary of Turkish-accented candidates increased when they looked German, $F(1,53) = 4.05$, $p = .049$, $\eta_p^2 = .07$, 95% CI [0.50, 195.33]. Again, as for competence and hirability, salary of

congruent Turkish candidates also increased, $F(1,53) = 6.94$, $p = .01$, $\eta_p^2 = .11$, 95% CI [29.92, 219.80].

Final evaluations. German-looking German-accented candidates were perceived as more hireable than all other candidates, $F_s \geq 11.38$, $p_s \leq .001$, $\eta_p^2_s \geq .17$ (other $F_s < 1$).

Sequence effects. The two sequences of presentation did not cause different evaluations $F_s \leq 3.24$, $p_s \geq .07$, $\eta_p^2_s \leq .03$.

Full description of Indirect Effects of Competence on Hirability

Table S3.

Correlations Between Dependent Variables. The Appearance First Experiment (1a) is Presented above the Line, The Accent First Experiment (1b) – below the Line.

1b	1a	Competence	Warmth	Hirability
Competence			.50***	.52***
Warmth		.57***		.43***
Hirability		.69***	.48***	

Note: *** $p < .001$

Appearance first. German-accented job candidates were perceived as more competent than Turkish-accented candidates, $b = .93$, $SE(\text{boot}) = .16$, $CI = [.62, 1.25]$ (Figure S3, upper panel). Appearance did not play a role, $b = -.16$, $SE(\text{boot}) = .16$, $CI = [-.47, .15]$, nor was there an interaction of appearance and accent, $b = -.13$, $SE(\text{boot}) = .22$, $CI = [-.57, .31]$. The more competent candidates were perceived, the more hireable they appeared, $b = .66$, $SE(\text{boot}) = .07$, $CI = [.52, .80]$. Disregarding competence, German-accented job candidates were perceived as more hireable, $b = .39$, $SE(\text{boot}) = .15$, $CI = [.10, .68]$. However, when including competence, Turkish-accented job candidates were perceived as more hireable, $b = -.63$, $SE(\text{boot}) = .19$, $CI = [-.99, -.26]$. Similarly, with competence included, Turkish-looking job candidates were perceived as more hireable, $b = -.58$, $SE(\text{boot}) = .18$, $CI = [-.92, -.23]$. The observed suppression effect suggests

that there are two contrary mechanisms, possibly higher competence of German-accented, but higher warmth of Turkish-accented candidates (see "negative suppression" in Conger, 1974; MacKinnon & Fairchild, 2009).

Furthermore, the interaction of appearance and accent on hirability was significant, $b = .88$, $SE(\text{boot}) = .25$, $CI = [.39, 1.36]$. For Turkish-looking candidates there was an effect of accent on hirability, $b = -.63$, $SE(\text{boot}) = .19$, $CI = [-.99, -.26]$, but for German-looking candidates there was not, $b = .25$, $SE(\text{boot}) = .18$, $CI = [-.11, .61]$. Indirect effects of accent on hirability through competence were significant for both, meaning that Turkish-looking candidates when speaking German were perceived as more competent and thus more hireable, $b = .62$, $SE(\text{boot}) = .12$, $CI = [.40, .87]$, and German-looking candidates when speaking German were also perceived as more competent and thus more hireable, $b = .53$, $SE(\text{boot}) = .13$, $CI = [.30, .80]$. The direct effect of accent was moderated by appearance, but the overall indirect effect was not, which was reflected by the non-significant moderated mediation index, $b = -.08$, $SE(\text{boot}) = .15$, $CI = [-.38, .19]$.³

³ For warmth neither of the indirect effects was significant (for Turkish-looking, $b = .07$, $SE(\text{boot}) = .08$, $CI = [-.10, .23]$; German-looking, $b = -.08$, $SE(\text{boot}) = .09$, $CI = [-.26, .08]$). Details are available from the first author.

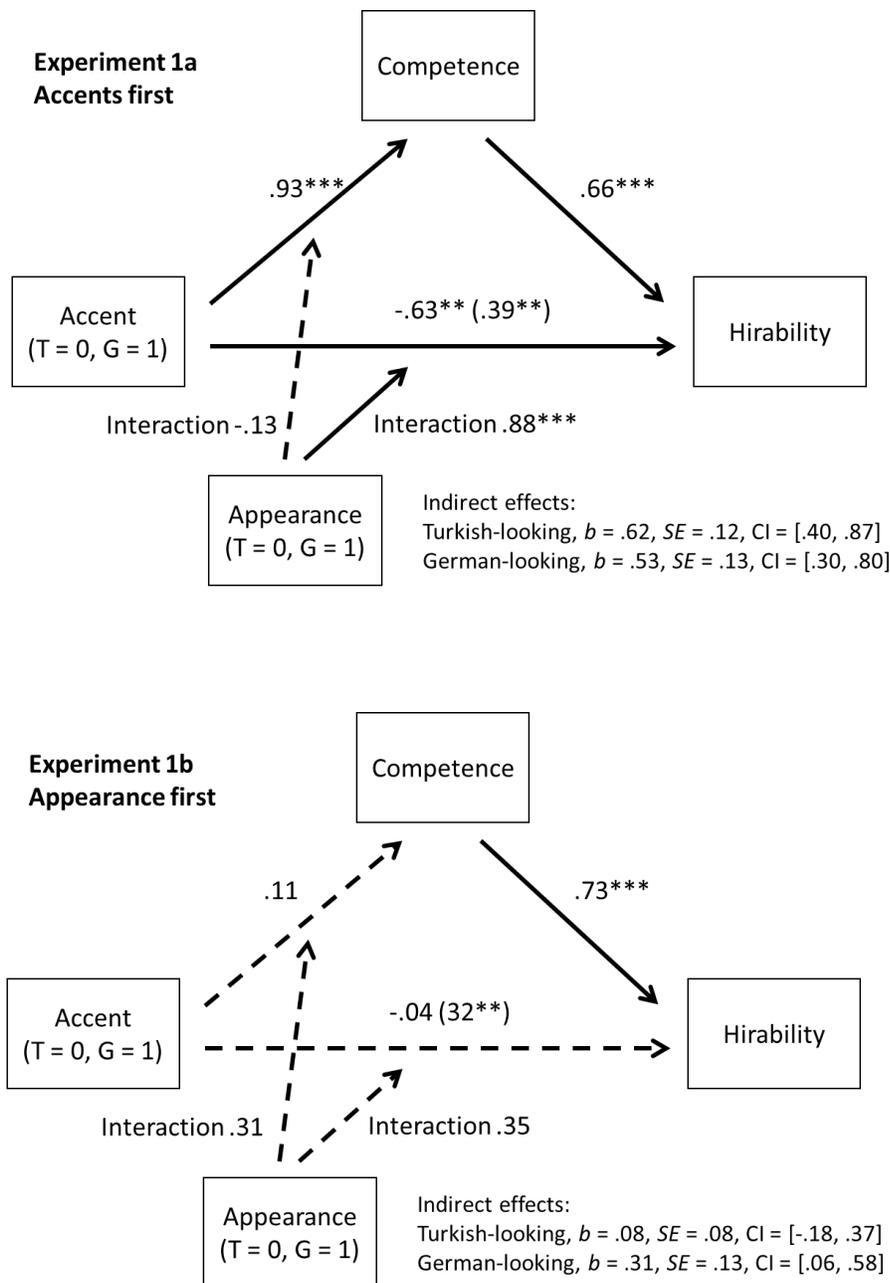


Figure S3. Indirect effects of accent and appearance on hirability via competence in Experiment 1a where accents were added to appearance (upper panel) and 1b where appearance was added to accents (lower panel).

Accent first. When accents were available from the beginning and appearance was added, neither accent, $b = .11$, $SE(\text{boot}) = .19$, $CI = [-.26, 0.47]$, nor appearance, $b = -.06$, $SE(\text{boot}) = .19$, $CI = [-.42, 0.31]$, nor their interaction, $b = .31$, $SE(\text{boot}) = .26$, $CI = [-.20, 0.83]$ influenced

competence evaluations (Figure S3, lower panel). Still, the more competent the candidates were perceived, the more hireable they appeared, $b = .73$, $SE(\text{boot}) = .06$, $CI = [.62, .84]$. Disregarding competence, German-accented job candidates were perceived as more hireable, $b = .32$, $SE(\text{boot}) = .14$, $CI = [.04, .61]$. When including competence, the effect disappeared, $b = -.04$, $SE(\text{boot}) = .15$, $CI = [-.34, .25]$. Appearance did not play a role, $b = -.01$, $SE(\text{boot}) = .15$, $CI = [-.31, .28]$.

Although the accent*appearance interaction was not significant, $b = .35$, $SE(\text{boot}) = .21$, $CI = [-.07, .76]$, results showed different direct effects of accent depending on appearance. These differences were opposite to the appearance-first sequence: for German-looking candidates, $b = .30$, $SE(\text{boot}) = .15$, $CI = [.01, .60]$, accent influenced hirability, but not for Turkish-looking candidates, $b = -.04$, $SE(\text{boot}) = .15$, $CI = [-.34, .25]$. The indirect effect of accent on hirability through competence was significant only for German appearance: for German-accented Turkish-looking candidates competence did not explain hirability, $b = .08$, $SE(\text{boot}) = .08$, $CI = [-.18, .37]$, but German-accented candidates when also German-looking were perceived as more competent and thus more hireable, $b = .31$, $SE(\text{boot}) = .13$, $CI = [.06, .58]$. Overall, for accents first, both the direct and indirect effects of accent were moderated by appearance, but these differences were small, reflected by a non-significant moderated mediation index: $b = .23$, $SE(\text{boot}) = .19$, $CI = [-.15, .60]$.⁴ In sum, presenting first candidates' faces versus voices influenced not only the final evaluations, but also partly the process of competence ascription and the following recommendations to hire this person.

⁴ For warmth neither of the indirect effects was significant (Turkish-looking, $b = -.13$, $SE(\text{boot}) = .09$, $CI = [-.32, .04]$; German-looking, $b = .15$, $SE(\text{boot}) = .10$, $CI = [-.04, .36]$).

References for the Appendix

- Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and psychological measurement, 34*(1), 35-46.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285-290. doi:10.1037/h0033731
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*, 1377-1388.
- MacKinnon, D., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science, 18*(1), 16-20.
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments & Computers, 36*(4), 630-633. doi:10.3758/BF03195583
- 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.
- Zuckerman, M., & Driver, R. E. (1989). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior, 13*(2), 67-82. doi:10.1007/BF00990791

All materials, datasets, results, and detailed procedures can be obtained from the authors via e-mail at karolina.hansen@psych.uw.edu.pl