Competent and Warm? How Mismatching Appearance and Accent Influence First Impressions

Karolina Hansen*

University of Warsaw

Tamara Rakić

Lancaster University

Melanie C. Steffens

University of Koblenz-Landau

*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

Funds

The current research was supported by the ProExzellenz program of the state of Thuringia, Polish National Science Centre (NCN grant DEC-2013/08/S/HS6/00573), Foundation for Polish Science (FNP START 030.2015-W scholarship) awarded to K.H., and by a grant of the German Research Foundation (DFG STE 938/10-2; FOR 1097) to M.C.S. and T.R.

Acknowledgements

We thank Claudia Niedlich for her help in data collection, Wiktor Soral for his help with converting the dataset, as well as Aysan Ashoee, Jana Meyer, the editor and two anonymous reviewers for their comments on a previous version of this manuscript.

Paper word count: 4970 (with footnotes, without tables)
Abstract

Most research on ethnicity has focused on visual cues. However, accents are strong social cues that can match or contradict visual cues. We examined understudied reactions to people whose one cue suggests one ethnicity, whereas the other cue contradicts it. In an experiment conducted in Germany, job candidates spoke with an accent either congruent or incongruent with their (German or Turkish) appearance. Based on ethnolinguistic identity theory, we predicted that accents would be strong cues for categorization and evaluation. Based on expectancy violations theory we expected that incongruent targets would be evaluated more extremely than congruent targets. Both predictions were confirmed: Accents strongly influenced perceptions and Turkish-looking German-accented targets were perceived as most competent of all targets (and additionally most warm). The findings show that bringing together visual and auditory information yields a more complete picture of the processes underlying impression formation.

Keywords: non-native speech, stereotypes, ethnolinguistic identity, expectancy violations, impression formation, person perception

Abstract word count: 142
Competent and Warm? How Mismatching Appearance and Accent Influence First Impressions

In today’s world of migration, people one meets have different cultural backgrounds (Davis, D’Odorico, Laio, & Ridolfi, 2013). National and ethnic distinctions in use for centuries are becoming outdated and inaccurate. As societies become more multicultural, we increasingly encounter people of mixed ethnicity, whose appearance and accent may violate expectations (King-O’Riain, Small, Mahtani, Song, & Spickard, 2014). In Germany, for instance, people may expect that a Turkish-looking person speaks German with a Turkish accent, and they may be surprised to hear native-like German (Hansen, Steffens, Rakić, & Wiese, in press). Up to now, social psychological research has largely overlooked the existence of such individuals and how impressions of them are formed. In the current research, we investigate how people evaluate others based on their appearance and accent, when one of these cues indicates that the person is an outgroup member (e.g., looks Turkish) and the other that the person is an ingroup member (e.g., speaks with a standard German accent).

When people encounter others, several cues indicating their ethnicity can be congruent or incongruent with each other. In the following, we focus on physical appearance (Dion, Berscheid, & Walster, 1972) and voice information (Zuckerman & Driver, 1989), specifically accent, as two powerful cues indicating social category memberships.

Language and manner of speaking are at the core of ethnolinguistic identity theory (ELIT; Giles, Bourhis, & Taylor, 1977; Giles & Johnson, 1981, 1987). Based on social identity theory (Tajfel, Billig, Bundy, & Flament, 1971; Tajfel & Turner, 1979), ELIT focuses on the importance of language and accent in identity formation and maintenance. ELIT postulates that language is the most important marker of ethnic identity and others’ impressions are often based on accents. Researchers have shown that people who speak with a nonstandard accent are perceived as less intelligent and of lower social status (Fuertes,
Gottdiener, Martin, Gilbert, & Giles, 2012; Giles & Powesland, 1975), but can also be seen as more loyal and sociable (Fuertes et al., 2012; Giles, 1971).

Accent-based discrimination is an unrecognized potential threat often overlooked in research and in real life (Crandall, Eshleman, & O’Brien, 2002; Hansen, Rakić, & Steffens, 2014; Ng, 2007). In the US 21% of the population speaks a language other than English at home and among them, 42% speak English less than very well (U.S. Census Bureau, 2011). In Germany about 9% of the population speaks a language other than German at home and 63% of them speak German less than very well (Haug, 2008). Thus, native speakers may expect that a foreign-looking person speaks with a foreign accent (Cheryan & Monin, 2005; Hirschfeld & Gelman, 1997).

Appearance and accent of a person can both be indicators of this person’s ethnicity. Therefore, people with mismatching appearance and accent could be difficult to categorize, others could be surprised when encountering them, and they could be evaluated differently than non-surprising people. There is very little research on reactions to people who embody conflicting cues about their categorization, such as mismatching appearance and accent.

To the best of our knowledge, only a few studies directly contrasted the role of appearance and accent in person perception. Two early studies did not aim at contrasting different types of cues, but found stronger effects of speech style than of racial labels on the perception of targets (Jussim, Coleman, & Lerch, 1987; McKirnan, Smith, & Hamayan, 1983). Later studies explicitly aimed at contrasting appearance and accent and showed that accent is a stronger cue than appearance for ethnic categorization in adults (Rakić, Steffens, & Mummendey, 2011), ingroup favoritism in children (Kinzler, Dupoux, & Spelke, 2007; Kinzler, Shutts, Dejesus, & Spelke, 2009), and beliefs about general knowledge of foreign-accented speakers (Rödin & Özcan, 2011). The effects were observed across cultures (US, France, Germany, Sweden) and with different dependent variables. Both Rakić and colleagues...
(2011) and Pietraszewski and Schwartz (2014b) independently ran similar who-said-what experiments in Germany and the US, to reveal that accent is crucial in social categorization.

Going beyond mere categorization, it is interesting how appearance-accent (mis)matches influence evaluations. A possible mechanism at work here could be expectancy violations. Expectancy violations theory postulates that violations of expectations produce more extreme outcomes than situations that match those expectations (e.g., Burgoon & Burgoon, 2001; Roese & Sherman, 2007). For example, Blacks with strong academic qualifications were evaluated as more competent than Whites with similar credentials, which represented positive violations of expectations based on the stereotype that Blacks are less academically-oriented (Jackson, Sullivan, & Hodge, 1993). Similarly, women with top credentials were evaluated more favorably as leaders than similarly qualified men because they violated stereotypical gender expectations (Rosette & Tost, 2010). Conversely, Whites who spoke nonstandard English were viewed more negatively than Blacks who did, representing negative expectancy violations (Jussim et al., 1987).

The Current Research

The present research examined how appearance and accent, suggesting the same or different ethnicities, influence how people are evaluated. We let our participants evaluate others on the two fundamental stereotype dimensions competence and warmth (Abele & Wojciszke, 2007; Fiske, Cuddy, Glick, & Xu, 2002). Because Turks are the largest immigrant group in Germany, we chose Germans and Turks as targets (Federal Ministry of the Interior, 2007). In Germany, as in the US and in many other countries, Turks (and Muslims more broadly) are stereotypically perceived as low on competence and warmth (Asbrock, 2010; Froehlich, Martiny, Deaux, & Mok, 2016). In contrast, the majority ingroup tends to self-stereotype as high on both dimensions (Fiske, Cuddy, & Glick, 2007; Fiske et al., 2002). In the case of Germans, in some studies they perceive themselves as competent and warm
(Asbrock, 2010; Eckes, 2002), in some as competent and moderately warm (Froehlich et al., 2016), and still in others as competent, but not warm (e.g., Cuddy et al., 2009).

In a computer-based experiment, we studied the influence of auditory and visual cues to ethnicity on the perceived competence and warmth of job candidates. We expected incongruent targets to violate participants’ expectations. Therefore, we also included a categorization task and tested whether incongruent targets were unexpected and thus categorized more slowly than congruent targets. We used photographs of male targets and recordings of speech in congruent or incongruent combinations. Male targets were used because stereotypes of nationalities apply more to men than women (Eagly & Kite, 1987) and for Germans the prototypical Turk is a man (e.g., Klingst & Drieschner, 2005).

Hypotheses 1-2 establish the basis for testing our main (evaluation) hypotheses. Based on ELIT (e.g., Giles & Johnson, 1987), we expected that accent would be a strong cue for social categorization. Conceptually replicating previous studies (Pietraszewski & Schwartz, 2014b; Rakić et al., 2011), in the German-Turkish context, targets should be categorized more according to their accent than appearance (Hypothesis 1a). We were especially interested in incongruent targets and we expected that Turkish-looking targets speaking standard German would be generally categorized as German (Hypothesis 1b) and German-looking targets with a Turkish accent would be categorized as non-German (Hypothesis 1c).

Research on expectancy-violations shows that counter-stereotypical people evoke more effortful cognitive processing than stereotypical people (Bettencourt, Dill, Greathouse, & Charlton, 1997; Roese & Sherman, 2007). When people meet a counter-stereotypical person, the discrepancy leads to re-categorization until an appropriate relevant category or subcategory is found (Fiske & Neuberg, 1990; Hutter & Crisp, 2006; Kunda & Thagard, 1996). Thus, we hypothesized that incongruent targets should be more difficult to categorize, which would be reflected by longer categorization reaction times (RTs, Hypothesis 2).
Based on ELIT, we predicted that accents would strongly influence evaluations (Hypothesis 3). Based on extensive research showing that nonstandard speakers are evaluated as less competent than standard speakers (Fuertes et al., 2012) and on the fact that Turks are perceived in Germany as incompetent (Asbrock, 2010; Froehlich et al., 2016), we expected Turkish-accented speakers to be evaluated as less competent than standard German speakers (Hypothesis 4). As findings regarding warmth of nonstandard speakers (Fuertes et al., 2012) as well as perceived warmth of Germans and Turks in Germany are mixed (e.g., Froehlich et al., 2016), we did not formulate specific predictions for this dimension.

Our main hypothesis was that incongruent targets would be evaluated differently than congruent targets. Based on expectancy violation theory (e.g., Burgoon & Burgoon, 2001; Roese & Sherman, 2007), we expected that incongruent targets would be evaluated more extremely than congruent targets in the direction of the valence of the violation. Again, as Turks are consistently perceived in Germany as incompetent and Germans as competent, but perceptions of their warmth differs between studies (Froehlich et al., 2016), we formulated these hypotheses for the competence dimension, but only explored the warmth dimension. Specifically, we hypothesized that when participants see a Turkish-looking person speaking standard German, their negative expectations would be positively violated and they would evaluate the target as very competent (Hypothesis 5a). Conversely, we expected that German-looking targets speaking with a Turkish accent would negatively violate participants’ expectations, and therefore be evaluated as incompetent (Hypothesis 5b).

**Method**

**Pre-Tests and Selection of Stimulus Materials**

We used portrait photographs of faces from an online database (Minear & Park, 2004) and added several of our own photographs of Turkish men. All men were young, with a neutral facial expression, without glasses, and with a neutral modern haircut. Pictures were converted into black and white.
Short voice samples of young German and Turkish native speakers were recorded. All speakers said the same neutral everyday phrase, “Good morning. Nice to meet you” (in German: “Guten Tag. Es freut mich, dass wir uns kennen lernen”), ensuring that any influence of the content of the statement was excluded and that accented sentences were easy to understand. Speakers were briefly trained, speech rate was held constant, and voice samples were approximately three seconds long.

To avoid the “what is beautiful is good” phenomenon (Dion et al., 1972; Zuckerman & Driver, 1989) and ensure that the stimuli were perceived as typical for their respective groups, all stimuli were pre-tested for attractiveness, pleasantness, ethnic typicality, and accent strength (Ryan, Carranza, & Moffie, 1977).

Pre-test participants ($N = 29$, 13 men, $M_{age} = 22.73$, $SD = 3.42$) were from the same population as participants in the main experiment, but participated only in the pre-test. The pre-test consisted of a block of faces and a block of voices. After each face or voice was presented in random order, participants answered how (1) attractive, (2) pleasant, (3) typically German, and (4) typically Turkish targets appeared or sounded ($1 = \text{not at all}$ to $7 = \text{very much}$). Voices were also evaluated regarding accent strength ($1 = \text{no accent at all}$ to $7 = \text{very strong accent}$).

From the pre-tested photographs of faces, we selected four moderately attractive and pleasant German and four Turkish-looking faces; all of them were typical for their respective groups (Table 1). Similarly, from the pre-tested voices, we selected four plus four moderately attractive and pleasant, but typical voices (Table 1).

---

1 One could worry that Turkish-looking faces were descriptively less attractive than German-looking faces. However, in the later evaluations Turkish-looking targets were perceived as most competent when they spoke standard German, but as least competent when they spoke with a Turkish accent. Thus, the descriptive difference in facial attractiveness of targets cannot account for the findings.
In the main experiment, we wanted to cross accents with appearance cues and study categorization and evaluation of mixed people. To be sure that accent and appearance had the same categorization baselines on their own, not yet being combined, we ran another pre-test ($N = 18$, $4$ men, $M_{\text{age}} = 26.06$, $SD = 6.31$). We used the same categorization task that we later used in the main study, but for the pre-test, we presented faces and voices separately (in two blocks with randomized block and stimulus order). The results showed that: German faces were in 100% of cases categorized as German, Turkish faces in 93% as Turkish, German voices in 90% as German, and Turkish voices in 93% as Turkish (Table 1). A $2 \times 2$ chi square test showed no differences between these percentages, $\chi^2 = 0.26$, $p = .61$.

**Table 1**

*Pre-tests Ratings of Photographs of Faces and Recordings of Voices*

<table>
<thead>
<tr>
<th></th>
<th>Faces</th>
<th></th>
<th>Voices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(\text{SD})_{\text{German}}$</td>
<td></td>
<td>$M(\text{SD})_{\text{Turkish}}$</td>
<td></td>
</tr>
<tr>
<td>Attractiveness</td>
<td>3.18 (1.21)</td>
<td></td>
<td>2.82 (1.04)</td>
<td></td>
</tr>
<tr>
<td>Pleasantness</td>
<td>4.47 (0.89)</td>
<td></td>
<td>4.14 (1.05)</td>
<td></td>
</tr>
<tr>
<td>Typically German</td>
<td>5.33 (1.29)</td>
<td></td>
<td>1.62 (0.70)</td>
<td></td>
</tr>
<tr>
<td>Typically Turkish</td>
<td>1.34 (0.47)</td>
<td></td>
<td>3.66 (1.71)</td>
<td></td>
</tr>
<tr>
<td>Accent strength</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Categorization</td>
<td>100%</td>
<td></td>
<td>93%</td>
<td></td>
</tr>
</tbody>
</table>

**Experimental Design**

The experiment had a $2$ (appearance: German vs. Turkish) $\times$ $2$ (accent: standard German vs. German with a Turkish accent) within-subject design. Thus, there were four target
types: German appearance/German accent (GG, congruent), Turkish appearance/Turkish accent (TT, congruent), German appearance/Turkish accent (GT, incongruent), and Turkish appearance/German accent (TG, incongruent). Stimulus composition was counterbalanced: any given voice (e.g., speaking standard German) was matched with a congruent picture (German-looking person) in one version of the experiment and with an incongruent picture (Turkish-looking person) in the second version. For generalization and control reasons, there were initially also two between-participants factors: context and face-voice sequence. As evaluations of others could depend on the context (e.g., Abele & Wojciszke, 2007; Vonk, 1999), we tested for the generalizability of our findings in the contexts of a roommate search and a job interview. The sequence of presentation was counterbalanced between participants: half of them first saw the face of a target and then immediately heard the voice, and half heard the voice and then immediately saw the face.

Participants

We stopped collecting data after achieving at least 50 participants per condition (context and sequence). Participants were 226 undergraduate students of various faculties of a German university. After excluding the data of 11 participants who were not native German speakers, the final sample consisted of 215 participants (72 men, $M_{age} = 22.33$, $SD = 3.24$). They were compensated with either €1 and a chocolate bar or with partial course credit.

Procedure and Measures

After being welcomed by an experimenter unaware of the study’s hypotheses, participants were seated in front of a computer screen and signed informed consent. The experiment consisted of an evaluation and a categorization block, with the same targets in each. First, participants were asked to imagine that either they were helping in a recruitment process at their workplace or that they had a free room for rent in their apartment (later analyzes showed no differences between these two contexts). All participants first saw two “filler” congruent German targets for training purposes and to set a common base. Then, the
main targets were presented in an individual random order. Targets’ faces and voices were presented with one second in between. For all targets, participants were asked to look at a face and listen to a voice and answer on a separate screen how competent (competent, competitive, independent, $\alpha = .93$) and warm (likeable, warm, good-natured, $\alpha = .91$) the person appeared ($1 = \text{not at all} \text{ to } 7 = \text{very much}$; Asbrock, 2010; Fiske et al., 2002). After this, participants saw and heard the same targets (in a different sequence) again and were asked to answer the question “Is this person German?” with yes and no as quickly as possible; RTs were collected. For categorization, we added female targets and questions about the gender of the target as filler items in order to prevent mental preparation to responding always to the same question, avoiding falsely short reaction times.² Finally, participants answered demographic questions, provided their email address for debriefing, were given their reward, thanked, and dismissed.

Results

Preliminary Analyses

Two preliminary 2 (appearance: German vs. Turkish) × 2 (accent: German vs. Turkish) × 2 (context: job interview vs. students’ apartment) mixed analyses of variance (ANOVA) yielded no effects involving context on competence or warmth evaluations (all $F$s < 1). Similar analyses including presentation sequence (appearance-accent vs. accent-appearance) yielded no main effects of sequence ($F$s < 1.94, $p$s = .17), and only one out of six possible interactions on the warmth dimension.³ Therefore, data were collapsed across these factors.

² A few supplementary questions (manipulation check: accent strength, cooperativeness, trustworthiness, suggested salary/room rent) yielded similar results but will not be reported for space concerns. Motivation to respond without prejudice was assessed at the end and did not moderate the findings.

³ An interaction of appearance, accent, and sequence on warmth evaluations, $F(1,197) = 18.93, p < .001$, $\eta^2 = .09$, boiled down to the following finding: German-looking Turkish-accented targets were perceived as warmer when their German appearance was presented first ($M = 4.91, SD = 1.58$) than
Social Categorization

Pre-requisites for analyzing the evaluations were that the data replicate the strong influence of accent on social categorization and that incongruent stimuli are expectancy-violating and thus people take longer to categorize them. As can be seen in Figure 1, targets were categorized more according to their accent than appearance, which was tested by means of a binomial logistic regression for repeated measures using the generalized estimating equations method (Zeger & Liang, 1986; see Table 2). As the Wald statistic shows, accent was a significant and strong predictor of categorization. The influence of appearance was much weaker and there was an interaction effect of appearance and accent. Follow-up analyses showed that German-looking targets were more often categorized as Germans than Turkish-looking targets, and this effect was stronger for German-accented speakers, McNemar’s $\chi^2 = 50.21, p < .001$, than for Turkish-accented speakers, McNemar’s $\chi^2 = 10.62, p = .001$, which could be due to a floor effect for Turkish-accented speakers. The results confirmed the Hypothesis 1a that accent would play a stronger role for categorization than appearance. Hypotheses 1b and 1c were also confirmed as Turkish-looking German-accented targets were mostly (65%) categorized as German and German-looking Turkish-accented targets as non-German (87%).

when their Turkish accent was first ($M = 4.56$, $SD = 1.63$), $F(1,197) = 4.76, p = .03, \eta^2 = .02$, but Turkish-looking German-accented targets were evaluated similarly in both presentation sequences, $F < 1$. This suggests that the sequence played only a minor and selective (or even random) role for evaluations.
Figure 1. Percent of targets categorized as Germans or non-Germans (left) and mean reaction times of categorization by target type (right). Error bars represent standard errors of the mean.

Table 2

Logistic Regression Results for Accent and Appearance Predicting Categorization of Targets as German or Non-German.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>B(SD)</th>
<th>95% CI</th>
<th>Wald</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.64</td>
<td>0.20</td>
<td>[-3.03, -2.24]</td>
<td>173.66</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Accent</td>
<td>4.54</td>
<td>0.24</td>
<td>[4.06, 5.02]</td>
<td>349.35</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Appearance</td>
<td>2.00</td>
<td>0.23</td>
<td>[1.55, 2.44]</td>
<td>77.25</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Accent*Appearance</td>
<td>-0.98</td>
<td>0.33</td>
<td>[-1.64, -0.33]</td>
<td>8.65</td>
<td>1</td>
<td>.003</td>
</tr>
</tbody>
</table>

Reaction times. We excluded responses that were ±3 standard deviations from the mean. We computed a 2 (accent: German vs. Turkish) × 2 (congruence of targets: congruent vs. incongruent) repeated measures ANOVA. The analysis showed that accent did not
influence RTs, $F < 1$, but congruence did, $F(1,197) = 7.61, p = .006, \eta_p^2 = .04$ (Figure 1).

Incongruent targets were categorized more slowly ($M = 1347.28$ ms, $SD = 539.57$ ms) than congruent targets ($M = 1250.58$ ms, $SD = 432.98$ ms), corroborating Hypothesis 2 that incongruent targets are more difficult to categorize (interaction: $F < 1$). Having confirmed that incongruent stimuli were expectancy-violating, we analyzed the effects of appearance and accent on evaluations.

### Competence Impressions

A 2 (appearance: German vs. Turkish) × 2 (accent: German vs. Turkish) repeated measures ANOVA showed that targets speaking standard German were evaluated as more competent ($M = 4.83$, $SD = 0.75$) than Turkish-accented targets ($M = 4.20$, $SD = 0.84$), $F(1,197) = 85.74, p < .001, \eta_p^2 = .30$ (Figure 2, Hypothesis 4). Competence evaluations also depended on appearance, but to a smaller extent, $F(1,197) = 6.21, p = .01, \eta_p^2 = .03$ (Hypothesis 3). More importantly, evaluations depended on specific combinations of accent and appearance, as reflected by an interaction effect, $F(1,197) = 20.63, p < .001, \eta_p^2 = .10$.

Analyses of simple main effects showed that among German-accented targets, Turkish-looking (i.e., incongruent) targets were perceived as more competent than German-looking (i.e., congruent) targets, $F(1,197) = 21.30, p < .001, \eta_p^2 = .10$ (Hypothesis 5a). Turkish-accented targets were evaluated as similarly competent whether they were German- or Turkish-looking, $F(1,197) = 2.47, p = .12, \eta_p^2 = .01$ (Hypothesis 5b).
Warmth Impressions

An ANOVA on the warmth dimension showed that neither accent itself, $F < 1$, nor appearance itself, $F(1,197) = 3.68, p = .06, \eta^2_p = .02$, influenced warmth evaluations in a significant way. Only the combination of appearance and accent influenced warmth perceptions, interaction: $F(1,197) = 38.52, p < .001, \eta^2_p = .16$. As depicted in Figure 2, incongruent targets were evaluated as warmer than congruent targets. More precisely, among German-accented targets, those who also looked German were perceived as less warm than those who looked Turkish, $F(1,197) = 31.38, p < .001, \eta^2_p = .14$; among Turkish-accented targets, those who also looked Turkish were perceived as less warm than those who looked German, $F(1,197) = 13.52, p < .001, \eta^2_p = .06$. These results show that incongruent targets were perceived as warmer than both congruent German and congruent Turkish targets.

Additional Analyses

In sum, Turkish-looking German-accented targets were evaluated as both most competent and, together with other incongruent targets, as most warm. German-looking
Turkish-accented targets were along with Turkish-Turkish targets evaluated as least competent. Overall, competence and warmth evaluations were correlated, $r = .66, p < .001$.

As targets were outgroup members for women, but gender-ingroup members for men, one may wonder whether effects differ between female and male participants. Exploratory analyses showed that the pattern of results for both genders was similar, but the differences between targets were larger for men. Especially the advantage of the Turkish-looking German-accented target over the German-German target was larger for men (both for competence and warmth). Considering the higher percentage (67%) of women in the sample, we conclude that the observed effects would be larger in a more balanced sample.

Studies show that it is crucial how a stereotype-incongruent person is categorized (e.g., Bless, Schwarz, Bodenhausen, & Thiel, 2001). To check potential influence of categorization on evaluation, we compared evaluations of the Turkish-looking German-accented targets between participants who categorized them as German (35%) or non-German (65%). Results showed no significant differences ($t < |1.7|$, $p > .10$ for competence, $t < |1|$ for warmth), indicating that categorization did not affect evaluations.

**Discussion**

When people encounter others, they often see and hear them. Their appearance and speech, as well as the combination of those two, can influence how people evaluate each other. Although such cross-modal effects are frequent in real life, they are relatively underrepresented in psychology (see also Freeman & Ambady, 2011; Zuckerman, Miyake, & Hodgins, 1991). The present research provides an original contribution to understanding the influence of visual and auditory cues on impression information. Targets were seen in photographs and heard in short voice recordings. They appeared Turkish or German and spoke standard German or German with a Turkish accent. Participants evaluated targets’ competence and warmth, and categorized them as Germans or non-Germans.
In a pre-test, appearance and accent presented separately were similarly used to infer ethnicity. When pitted against each other, accent was more diagnostic for social categorization and evaluation. Such a strong role of accent is in line with ethnolinguistic identity theory (Giles & Johnson, 1987) and results of research conducted in the US (Kinzler et al., 2009; Pietraszewski & Schwartz, 2014b), Germany (Rakić et al., 2011), and Sweden (Rödin & Özcan, 2011). Nevertheless, it is an open question whether this would replicate everywhere or would depend on the diagnosticity of accents and appearance in a specific cultural context (see Pietraszewski & Schwartz, 2014a). Future cross-cultural research or experimental manipulations of diagnosticity could shed more light on this issue.

Our results also showed that standard German speakers were overall evaluated as more competent than Turkish-accented speakers. However, the evaluation of targets depended on the combination of their appearance and accent. As expectancy violations theory predicted (Burgoon & Burgoon, 2001), effects of appearance-accent mismatch went beyond “the sum of the elements” and Turkish-looking German-accented targets were perceived as most competent. German-looking Turkish-accented targets were, together with congruent Turkish targets, evaluated as the least competent. Thus, our expectancy violations hypotheses were confirmed.

As earlier findings about the perceived warmth of Turks in Germany (e.g., Froehlich et al., 2016) and of foreign-accented speakers (Fuertes et al., 2012) were inconclusive, we did not formulate hypotheses for warmth. Nonetheless, results on this dimension were very interesting: the two types of incongruent targets were evaluated as warmer than the two types of congruent targets. Congruent Turkish targets were perceived as relatively cold (see also Asbrock, 2010; Eckes, 2002), but targets who had only one Turkish feature were evaluated as warmer. Possibly, perceived threat changes evaluations (Cottrell & Neuberg, 2005): A young Turkish-looking Turkish-accented man might be too threatening to be perceived as nice and...
friendly, but when he possesses only one Turkish trait, stereotypes about warm Turks might be activated and expressed.

An alternative explanation pertains both to the warmth and the competence findings: According to expectancy violations theory or to the ‘black sheep’ effect (Biernat, Vescio, & Billings, 1999), German-looking Turkish-accented targets should be evaluated as least competent, but they were evaluated as similarly (in)competent as the congruent Turkish targets. This result suggests that other cognitive processes could also contribute to the observed effects. It could be that (appearance or accent) cues change their meaning in the context of other cues (e.g., Anderson, 1971; Kunda & Thagard, 1996). In a study where participants indicated how they interpreted surprising combinations of appearance and accents, German-looking faces sometimes changed the perception of Turkish accents: Some participants re-interpreted the targets as Northern or Eastern Europeans (Hansen, 2013). This shows how surprising combinations of accent and appearance can strongly change people’s perceptions (Kunda & Thagard, 1996; Remedios, Chasteen, Rule, & Plaks, 2011).

Turkish-looking but German-accented targets were evaluated as both warmest and most competent. The other incongruent targets, German-looking but speaking with a Turkish accent, were perceived as low in competence, but high in warmth. While the latter result could be interpreted as compensatory stereotyping (Yzerbyt, Provost, & Corneille, 2005), the earlier one could not. More research is needed to better understand the obtained findings on the warmth dimension and generally, the inconsistent findings for warmth evaluations in research on accents (Fuertes et al., 2012).

The positive evaluation of Turkish-looking targets who spoke standard German was in line with expectancy violations: Participants were positively surprised by these targets (which was reflected in longer categorization latencies) and evaluated them extremely well (Burgoon, 2009; Roese & Sherman, 2007). Previous research has shown, for example, that Blacks with strong academic qualifications were evaluated as more competent than Whites with similar...
credentials (Jackson et al., 1993). Along with other results interpreted as expectancy violations, these results can be also seen as an effect of lower linguistic standards that Germans might have for foreign-looking people. People may be evaluated in comparison to the average of their group (Biernat & Manis, 1994). Stereotype-incongruent targets can be contrasted from the group norm and described in such terms, for example: “For a Turk he speaks German very well” (Collins, Biernat, & Eidelman, 2009). Similar contrastive judgment patterns can also occur when expectations are violated. The present experiment offers no direct evidence of expectancy violation. However, in cases in which the same result can be based on different cognitive processes, measuring its neural correlates can be a useful tool for constraining explanations of such behavioral data (Bartholow, 2010). Research related to the present study, combining accent and appearance ethnicity cues, has shown that incongruent targets evoke brain reactions that can be interpreted as expectancy violations (Hansen et al., in press).

The results of studies like the present one may depend on the cultural context where they are conducted, for example, in a traditionally monocultural or multicultural country. The results could also depend on the characteristics and beliefs about a specific ethnic group. We chose Turks as targets as they are the biggest and the most prototypical immigrant group in Germany (Federal Ministry of the Interior, 2007). We are not aware of any data directly showing relationships between Turkish appearance, accent strength, and other variables. However, existing data show that about 7% of the German population speak German less than very well (Haug, 2008) and children of Turkish origin often have problems at school because of their low German language competencies (e.g., Becker, 2010). Thus, indirectly it can be inferred that Turkish appearance and Turkish accent are significantly related with each other in Germany, and also that Turkish accent could be related to low perceived acculturation. We cannot generalize our findings, for example, to Asians, who are often perceived to be
particularly competent and hardworking, and people could hold different assimilation expectations regarding them (Asbrock, 2010; Fiske et al., 2002).

Similarly, we had few and pre-tested stimuli per condition, which made the experiment well-controlled and its interpretations cleaner. The chosen stimuli were judged as typical for their groups, but we cannot know how representative they were of the Turkish- and German-origin populations in Germany. Nevertheless, we think that even if a specific cultural context or stimulus choice may have influenced the results, the mechanism is still interesting: If people expect from a Turkish-/Moroccan-/Indian- or German-/French-/American-looking person to speak with a specific accent, but the person speaks with a different one, this can be surprising, new qualities can emerge from such atypical combinations of features, and they can strongly influence evaluations.

Our results suggest that Turks in Germany would benefit from learning German at an early age, as foreign-looking people who speak standard German evoke an especially positive impression. We think that these are reasonable conclusions, but we would also like to draw attention to the other side of the coin. A widespread approach to communication problems between native and nonnative speakers is to reduce the accent of the nonnative speaker (e.g., Carlson & McHenry, 2006; Shah, 2012). This focuses attention only on one person’s responsibility, and eradicating accent in speech is very difficult or even impossible to achieve (Gluszek & Dovidio, 2010). We believe that to diminish language-based discrimination, it is important to address the role of native speakers’ consciousness, for example by using perspective-taking interventions (Hansen et al., 2014; Subtirelu & Lindemann, 2014).

Conclusions

An important implication of the present research is that researchers should pay more attention to the interactions of appearance, accent, and other cues in impression formation. Reactions to people with features suggesting different ethnicities have been little studied, but with increasing global mobility they are timely and important. With our research, we hope to
pave the way for future research on the social perception of people whose social categorization is ambiguous.
References


doi:10.1006/jesp.1996.1323


doi:10.1037/0022-3514.53.3.451


Froehlich, L., Martiny, S. E., Deaux, K., & Mok, S. Y. (2016). “It’s their responsibility, not ours”: Stereotypes about competence and causal attributions for immigrants’ academic underperformance. *Social psychology, 47*(2), 74-87.


