Journal of Personality and Social Psychology

# Blinded by the Accent! The Minor Role of Looks in Ethnic Categorization

Tamara Rakić, Melanie C. Steffens, and Amélie Mummendey

Online First Publication, November 1, 2010. doi: 10.1037/a0021522

Journal of Personality and Social Psychology

0022-3514/10/$12.00 DOI: 10.1037/a0021522

# Blinded by the Accent! The Minor Role of Looks in Ethnic Categorization

Tamara Rakić, Melanie C. Steffens, and Amélie Mummendey
Friedrich Schiller University Jena, Germany

The categories that social targets belong to are often activated automatically. Most studies investigating social categorization have used visual stimuli or verbal labels, whereas ethnolinguistic identity theory posits that language is an essential dimension of ethnic identity. Language should therefore be used for social categorization. In 2 experiments, using the "Who Said What?" paradigm, the authors investigated social categorization by using accents (auditory stimuli) and looks (visual stimuli) to indicate ethnicity, either separately or in combination. Given either looks or accents only, the authors demonstrated that ethnic categorization can be based on accents, and the authors found a similar degree of ethnic categorization by accents and looks. When ethnic cues of looks and accents were combined by creating cross categories, there was a clear predominance of accents as meaningful cues for categorization, as shown in the respective parameters of a multinomial model. The present findings are discussed with regard to the generalizability of findings using one channel of presentation (e.g., visual) and the asymmetry found with different presentation channels for the category ethnicity.

*Keywords:* accents, social categorization, "Who said what?" paradigm, multinomial model, ethnolinguistic identity theory

People we meet in our everyday life are extremely complex and multifaceted. Each individual belongs to multiple social categories simultaneously. The same category can even be present in more than one cue. Language, as one possible cue, can influence perception and impression formation in many ways, starting from what one has said to how it was said. For instance, when thinking about the population of the European Union, one grasps quickly that the visual differences between members of its constituent countries are not very big in contrast to the many languages that are spoken. Imagine an Italian and a German woman interacting. On the basis of their looks alone, one can imagine that the given difference would not be as clear as in the case of, for example, a man and a women interacting. However, as soon as they start speaking German, one would "see" the difference between them.

How and when people categorize others is one of the big questions in social psychology where the catch-22 between experimental control and ecological validity becomes obvious. The vast majority of studies have only used visual stimuli (including group labels) to indicate different social categories, even with regard to those categories in which other types of stimuli could be as important, or more important. The aim of the present article was to test social categorization when ethnicity, as a category, is presented with either looks or accents, or both.

## Social Categorization

Categorization is natural and an inevitable tendency of the human mind (Allport, 1954). There has been a wide range of research on categorical thinking (e.g., Brewer, 1988; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000; Sherman, Castelli, & Hamilton, 2002). It seems that a need for better understanding, as well as a need to simplify the complexity of the world, lead people to spontaneously use categories (for reviews on simple and cross-categorization, see Crisp & Hewstone, 2006; van Knippenberg, van Twuyver, & Pepels, 1994). Similar rules are applied both to physical objects and to people.

A fundamental characteristic of person perception and categorization is that people react to the first available and meaningful information to categorize others (Fiske & Neuberg, 1990). Features that are perceived as less informative are also less likely to be used in impression formation or categorization. Hence, the number of all possible categories that could potentially be applied is limited to those that are highly informative, whereas at the same time, less informative categories are being inhibited (cf. Stangor, Lynch, Duan, & Glas, 1992). In other words, the way one perceives or categorizes a person depends on the selection of information that is available and its salience or importance. Theoretically, category salience is a function of accessibility and fit. In other words, categories become salient when they fit the data and are accessible at the same time (Bruner, 1957; Oakes, 1987). *Accessibility* refers to the simplicity with which a category comes

Tamara Rakić, DFG Research Unit-Person Perception, Institut für Psychologie, Friedrich-Schiller-Universität Jena, Germany; Sozialpsychologie, Friedrich-Schiller-Universität Jena, Germany; Melanie C. Steffens, Institut für Psychologie, Friedrich-Schiller-Universität Jena; Amélie Mummendey, Institut für Psychologie, Lehrstuhl für Sozialpsychologie, Friedrich-Schiller-Universität Jena.

The present research was supported by the 3-year funding of Tamara Rakić as a doctoral student in the International Research Training Group (GRK622) "Conflict and Cooperation between Social Groups" granted by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and by DFG Grant Ste 938/10-1 awarded to Melanie C. Steffens. We thank Karolina Hansen, Friederike Feuchte, and Alison Benbow for valuable comments on an earlier draft of this paper.

Correspondence concerning this article should be addressed to Tamara Rakić, DFG Research Unit-Person Perception, Friedrich-Schiller-Universität Jena, Leutragraben 1 D-07743 Jena. E-mail: tamara.rakic@uni-jena.de

to mind, or, more specifically, it refers to the probability that a category will be used given appropriate external stimuli (e.g., Bruner, 1957). The chronic accessibility of categories such as gender, age, and race or ethnicity has been proposed to explain their prominence (e.g., Stangor et al., 1992).

Although verbal category labels have been used in many studies that have yielded important insights concerning the categorization process, mere verbal labels cannot grasp the complexity of the social world (Macrae & Bodenhausen, 2000). Indeed, verbal stimuli represent individual categories, while at the same time they neglect all other information that is present in real-life encounters. In other words, category labels polarize one's perception (Yamauchi & Yu, 2008). Gilbert and Hixon (1991) used the example of a "Black fireman" to explain how, for literate adults, reading a multiple label would simultaneously activate both categories. However, in the case of actually encountering a Black fireman, it could be that only fireman is activated, because it is more salient or more appropriate from the context. Empirical evidence indicates that people require different information-processing solutions when confronted with labels and people (Gilbert & Hixon, 1991). Zàrate and Smith (1990), while commenting on the use of the label *Black person* noted that: "Linguistic descriptions . . . force a single categorization, in contrast to the real person who is not only black but (perhaps) young, male, well dressed, tall, *speaking with Southern accent*, and so on" [italics added] (p. 162).

A question that arises is whether the same category would be perceived as equally meaningful when presented visually or auditorily. In fact, one of most fascinating aspects of human perception is the extraordinary ability to integrate different sensory modalities (i.e., vision, hearing, touch, taste, smell). Undeniably, seeing a visual representation of a given target while hearing the target speak is closer to a real-life situation of encountering others than a verbal label alone and appears very informative in (research on) person perception (cf. Kamachi, Hill, Lander, & Vatikiotis-Bateson, 2003; see also Zuckerman, Miyake, & Hodgins, 1991). Though auditory stimuli have been used in some previous studies (e.g., Mummendey, Otten, Berger, & Kessler, 2000; S. E. Taylor, Fiske, Etcoff, & Ruderman, 1978), they were almost exclusively used as a carrier of content (i.e., of stereotypical statements). One exception to this practice was made by Yzerbyt, Leyens, and Bellour (1995) in a study on the ingroup overexclusion effect. The authors used Belgium as a bilingual context and native French (i.e., Wallon) and Dutch speakers (i.e., Flemish) to indicate group membership of target speakers. The participants, who themselves were members of one of the two groups, had to decide whether a given target speaker was an in- or outgroup member. The results confirmed the ingroup overexclusion effect especially when short outgroup sentences were spoken by an ingroup member, because that made spotting the accent more difficult: In doubt, participants preferred to exclude a given target from the ingroup. More generally, these results show that accents are sometimes used for ingroup–outgroup categorization.

## Language and Ethnicity

Spoken language (i.e., speech) is content itself and not just a simple carrier of content (Giles & Powesland, 1975; see also Giles & Billings, 2004, for a 40-year review). In ancient history, language was already seen as an important part of every single person, as is testified by the origin of the word *personality* that comes from Latin *per* (through) and *sonare* (to sound), hence the literal meaning is "to sound through." Empirical studies confirm people's tendency to use speech style and voice as relevant information for evaluating personality traits of others (e.g., Allport & Cantril, 1934; Zuckerman & Driver, 1989; Zuckerman et al., 1991). Ko, Judd, and Blair (2006; see also Ko, Judd, & Stapel, 2009) recently showed that the quality of a voice (i.e., femininity) had a stronger influence on stereotype application than target gender itself. Additionally, different manners of speaking the same language are associated with different stereotypes, as in the example of the French language in France, Belgium, and Switzerland (Yzerbyt, Provost, & Corneille, 2005).

Ethnolinguistic identity theory (ELIT; Giles, Bourhis, & Taylor, 1977; Giles & Coupland, 1991; Giles & Johnson, 1981, 1987) indicates the importance of language for the ethnic categorization of self and others. This theory is based on communication accommodation theory (CAT; Shepard, Giles, & Le Poire, 2001; see also Giles, 1973, for the original speech accommodation theory). CAT attempts to explain possible reasons and effects of changes in manner of speaking in different communication situations. In any given interaction, one may or may not adjust one's way of speaking based on one's interlocutor and context. One can have different reasons for not accommodating. For example, nonaccommodation can be used as a tool to distinguish oneself from others or due to incapacity to converge to standard language. In addition, different attributions can be made by listeners for a lack of accommodation. Nonaccommodation (i.e., a nonstandard accent) due to a lack of ability to accommodate can involuntarily reveal category membership (e.g., ethnic background). This lack of accommodation is very often the case with different ethnic minorities in a host country (Bourhis, Giles, & Tajfel, 1973). The stronger the degree of a nonstandard accent, the more negative are the evaluations of targets (e.g., Ryan, Carranza, & Moffie, 1977). Considerable evidence shows that (degree of) nonstandard accent serves as a cue in the evaluation of speakers and influences their likelihood of getting a job (e.g., Aboud, Clement, & Taylor, 1974; Elwell, Brown, & Rutter, 1984). There is also nonaccommodation within the same language (e.g., dialectic diversity). Moreover, the mere presence of an accent is enough to indicate ethnic background. "Ethnicity is a slippery concept" (Gudykunst & Ting-Toomey, 1990, p. 310), but at the same time there is strong evidence for a language-ethnicity bond. Indeed, ethnic categorization is often manifest in the distinctiveness of different languages or dialects (or even accents). Even small children prefer same-accent friends independently of their race, indicating the centrality of accents (i.e., language) in categorizing others (e.g., Kinzler, Dupoux, & Spelke, 2007; Kinzler, Shutts, Dejesus, & Spelke, 2009). Additionally, language has been indisputably acknowledged to be one of the most essential dimensions of ethnicity (D. M. Taylor, Bassili, & Aboud, 1973). Changes of ethnic identity are accompanied by changes in language preference and attitude (e.g., McNamara, 1987). Given this central importance of language for ethnicity, one would expect to find that studies on ethnic categorization have used language variations besides visual stimuli (e.g., faces, names, etc.). This is not the case (cf. Ko et al., 2006, for similar conclusions; see also Gluszek & Dovidio, 2010, for a recent review on the stigma of accents).

In a nutshell, traditional lines of research (using almost exclusively visual stimuli or labels) may not offer a complete picture about social categorization that accounts for the full complexity of reality. Our main research question was to test how looks and accents determine ethnic categorization, both separately and in combination. Even though we are not actually using different categories theoretically, as is the case in cross-categorization studies, two different outcomes are possible when both looks and accents provide ethnic information (e.g., targets look typically Italian, or not, and they speak with an Italian accent, or not). First, it is possible that both sources of information have an additive relation and influence one another in categorization (e.g., Stangor et al., 1992; Zuckerman et al., 1991). This would mean that both sources of information are simultaneously used to classify a given target, resulting in the most informative outcome (i.e., categorization on the "subgroup" level). The second option is that due to limited cognitive capacity, inhibition of one source of information may occur. In other words, if there is an accentuation based on accents (e.g., an Italian accent), then different looks (Italian and German) could be assimilated (cf. Klauer, Ehrenberg, & Wegener, 2003). The accentuation hypothesis may appear unlikely, given that we are not using different categories, but different sources of information on the same categories. In other words, inhibition of information about a typical Italian looks in the presence of a typical Italian accent appears far-fetched. However, this language-based accentuation hypothesis is in line with ELIT's focus on language.

## Overview of the Experiments

In the European context, visual differences between people from different countries are clearly perceivable. Most Europeans would claim to recognize the typical Italian, typical German, or typical British looks. We used Italian versus German targets in our experiments with German participants because Italians are not uncommon, yet they are not strongly associated with negative stereotypes or lower status (cf. Wenzel, Mummendey, & Waldzus, 2007). There are some studies covering stereotypes related to Germans and Italians (e.g., Giles & Niedzielski, 1998; Krueger, 1996), though to the best of our knowledge there are no studies looking systematically at German stereotypes of Italians (see also Bianchi, Mummendey, Steffens, & Yzerbyt, 2010, for some findings on Italian stereotypes as compared with Germans). Still it seems that there is general stereotype acceptance among different nationalities.

The first experiment set the stage for the main research question. We separately tested the use of looks and accent to indicate ethnicity. In one condition, ethnicity[1] was indicated with typical German and Italian looks only, and in the other with accents only (standard German vs. German with Italian accent). To anticipate, looks and accent, when present separately, provoked similar degrees of ethnic categorization. In Experiment 2, we then tested ethnic categorization when both looks and accents in cross-categories were used (e.g., there were some typically German-looking targets, speaking with Italian accents).

In both experiments, we used the "Who Said What?" paradigm (WSW; S. E. Taylor et al., 1978). In this paradigm, participants are asked to observe a discussion of a small group. Subsequently, they are asked to match speakers and statements. The rationale is: If participants use a social category to organize information in memory, then more within-category confusions (e.g., confusing two Italian targets with one another) than between-category confusions (e.g., confusing an Italian target with a German one) should occur. Thus, the spontaneous activation of social categories can be investigated because participants are not explicitly asked to categorize targets (for an overview of WSW studies, see Klauer & Wegener, 1998). Previous studies using the WSW paradigm to investigate race or ethnicity (e.g., Frable & Bem, 1985; Hewstone, Hantzi, & Johnston, 1991; Kurzban, Tooby, & Cosmides, 2001; Stangor et al., 1992; S. E. Taylor & Falcone, 1982) have not made use of speech as a category indicator.

### Material Selection

As stimulus sampling is a crucial starting point of any research (Wells & Windschitl, 1999), we preselected and subsequently pretested all faces and voices we used. To assure that the stimuli used in our studies were adequate, we performed several independent tests with participants who did not participate in any of our experiments.

People can differentiate between several degrees of accentedness (Ryan et al., 1977). In order to describe our targets with regard to this feature, we conducted a test with $N = 30$ students ($M_{\text{age}} = 21.20$, $SD = 2.58$) to check for accentedness and fluency of each speaker on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very strong accent/very fluent*). As expected, Italian speakers were perceived as having a very strong accent ($Ms = 6.07$) compared with German speakers ($Ms = 2.07$), $F(1, 29) = 863.26$, $p < .05$, $\eta_p^2 = .97$. The opposite pattern was found for perceived fluency, $F(1, 29) = 151.55$, $p < .05$, $\eta_p^2 = .84$, with German speakers ($Ms = 6.18$) being perceived as more fluent compared with Italian speakers ($Ms = 3.74$). However, because the perceived fluency of Italian speakers was not different from the scale midpoint, $t(29) = -1.53$, $p = .14$, this did not negatively influence comprehensibility.

For the purpose of comparing the influence of looks and accents on categorization, it was vital to test the prototypicality of our targets for the respective categories. A total of 42 students from the Friedrich Schiller University Jena, Germany (30 women and 12 men, $M_{\text{age}} = 22.95$) evaluated all faces and voices in an analogous way. The order of the face and voice evaluation was counterbalanced. In addition to prototypicality, attractiveness was rated where large variations among stimuli might influence findings.

Faces of target males were portrayed in black-and-white photographs, with a respective height and width of $400 \times 528$ pixels on a screen resolution of $1024 \times 768$ pixels. We decided to use only male targets to avoid any possible gender influence on our findings (e.g., Lambert, 1967). All target voices and faces were rated for attractiveness as well as for both German and Italian prototypicality on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*). We thus obtained four typical Italian targets ($Ms = 5.4$ for prototypical Italian and 1.3 for prototypical German), four typical German targets ($Ms = 1.77$ and 5.82, respectively), and eight neutral targets ($Ms = 3.47$ and 4.24, difference: $t < 1$). All three groups of targets were equal on attractiveness (all $t$s $< 1$).

[1] We use the general term *ethnicity* even though we are aware that throughout history, there was mixing of ethnicities in Italy and Germany.

Speakers were native German and Italian men who were trained to speak the statements fluently with the same speed in order to avoid a possible influence of speech rate. All speakers spoke flawless German (i.e., without any grammatical errors); the only difference was the presence or absence of an Italian accent. The native German speakers each spoke without a distinctive or discernible dialect (we refer to them as "standard German speakers"), whereas the content of the statements (same also for native Italian speakers) was not in any way too formal or unusual for university context because all statements have been already extensively pretested (see Klauer & Wegener, 1998). All native Italian speakers were from northern Italy so that individual differences within Italian accents would not play any role (even though it is reasonable to assume that German participants would not be sensitive enough to perceive these differences). In addition to ratings of voice attractiveness and German and Italian prototypicality of accents, participants also judged whether speakers with an Italian accent were easily comprehensible or demanding to follow, and because no problems were reported ($M_{comprehensible}$ = 6.17, $M_{demanding}$ = 1.88), all speakers were used. We thus obtained a sample of six native German (*Ms* = 5.69 for prototypical German and 1.43 for prototypical Italian) and four native Italian speakers (*Ms* = 6.16 for prototypical Italian and 1.82 for prototypical German) who were judged as equally attractive ($t < 1$). A comparison of the face and accent ratings on prototypicality yielded no evidence for the potential assumption that differences between categories on the face dimension were smaller than those on the accent dimension ($t < 1$).

Finally, a test with 18 students (11 women and seven men; $M_{\text{age}}$ = 23.39) was run to test the perceived similarity of different voices within the same accent group. This test was run in order to exclude that the similarity among the Italian voices exceeded that among the German voices. Participants heard different speakers one after the other and had to compare them pairwise. The comparison was done on a 7-point Likert scale, with 1 indicating *definitely a different speaker*, 4 *not sure*, and 7 *definitely the same speaker*. We counterbalanced the order of voice presentation between participants (no significant interaction was found for order or participants' gender, all $F$s $< 1$). We found a significant effect of speaker's accent group on perceived similarity, $F(1, 14) = 8.13$, $p < .05$, $\eta_p^2 = .37$, with German native speakers being perceived as more similar ($M = 3.81$, $SD = 1.21$) than Italian native speakers ($M = 3.25$, $SD = 0.96$). A $t$ test confirmed that only Italian voices were significantly different from the midpoint of the scale, $t(17) = -3.33$, $p < .05$, compared with German voices ($t < 1$). In other words, people were unable to tell whether two German speakers were the same person or not, whereas they could distinguish the Italian speakers as different people. More important, the direction of this effect is opposite to an outgroup homogeneity effect (e.g., Ackerman et al., 2006; S. E. Taylor et al., 1978). In this way, it was possible to attribute any possible finding of outgroup homogeneity in the categorization experiments to accent group and not to the materials used.

With two additional preexperiments (total $N = 144$), we established that ethnic categorization in the original WSW paradigm is triggered both by the visual stimuli (typical German vs. Italian looks) and the auditory stimuli we used (standard German vs. German with Italian accent). The same effects were found whether the other stimuli were present or absent (i.e., in visually cued categorization, the voice was either present, but insignificant, or completely absent, and in auditory-cued categorization, neutral faces were present or completely absent). We replicated these preexperiments in Experiment 1 by using a multinomial modeling approach.

## Experiment 1

The traditional WSW paradigm (S. E. Taylor et al., 1978), as used in our preexperiments, has been criticized by Klauer and Wegener (1998) in several respects. First, participants are forced to guess whether they do not remember a given statement. Participants frequently fail to distinguish between old and new items, thus their item memory is less than perfect. To the degree that participants are only guessing, between and within category errors level themselves out. This is even more probable when there is no structural fit (Oakes, 1987): if the content of a statement is useless for identifying the source (as is the case in our experiments). Second, the same outcome (e.g., a correct assignment of a statement to a speaker) could be caused by different underlying processes (i.e., memory vs. guessing) that cannot be discriminated in the traditional WSW paradigm. A third problem concerns person memory. If one particular target is so well remembered that almost all of his statements are correctly assigned to him, the difference of within- and between-category errors would be zero, even though his category membership might be highly salient. Most important, the multinomial model allows comparing the activation of a target's social category (on the basis of looks vs. accents) independent of remembering the individual target and independent of guessing processes. Overall, misinterpretations of results obtained by means of the original WSW paradigm can be avoided using a multinomial model.

It would be useful to discriminate between item memory (memory for a given statement); person memory (memory that a specific speaker made a given statement); category memory (given item discrimination, but a lack of person memory, is there memory for a target's category membership?); and the respective guessing processes. These different parameters are of central importance, especially the category memory parameters, as they can show whether two stimulus types used (i.e., looks and accents) are comparable or different to begin with (see Klauer & Wegener, 1998, Experiment 3 for details). By making a minor change to the original WSW procedure, these cognitive processes can be estimated using a multinomial processing tree model (Riefer & Batchelder, 1988) introduced by Klauer and Wegener: In the matching task, besides statements from the discussion part, participants also receive new statements. Only if a given statement is categorized as old (i.e., previously presented), all the targets appear, and the participant indicates which target made this statement.

The main aim of Experiment 1 was to test whether ethnic categorization would be observed to the same degree in both the face (i.e., typical looks) and voice (i.e., accents) condition. Whereas finding ethnic categorization in the face condition would be a conceptual replication of previous findings, ethnic categorization based on accents has not been demonstrated before.

### Method

**Participants.** A total of 60 students from the Friedrich Schiller University Jena participated. Half of them were randomly assigned to the *Face* condition (22 women and 5 men), the other half to the *Voice* condition (19 women and 10 men). The remaining four participants did not indicate their gender. The participants' age ranged from 18 to 31 ($M$ = 21.78, $SD$ = 2.71).

After giving their consent to take part in the experiment, participants undertook the experiment individually on the computer, in one session of about 20–25 min. At the end, participants were fully debriefed and rewarded for their participation with either a chocolate bar or partial course credit.

**Materials and procedure.** In the Face condition, the pretested photographs of four Italian- and four German-looking targets were used, while the statements they allegedly made were presented in written form. Similarly, in the Voice condition, eight neutral looking male targets, four of them paired with statements in standard German and four with statements spoken with an Italian accent by the men whose voices had been pretested were used. Statements from a pool of around 200 statements used in previous studies on social categorization (Klauer & Wegener, 1998) were used.[2] Statements expressed general critical points about the university on various particular topics, for instance, exams, seminar rooms, assignments, and the library. From this pool, a total of 48 statements were used in the discussion (another 48 statements on the same topics were presented only in the matching part). Statements were randomized in such a way between targets that each made six statements on six different subjects. For instance, every target made one statement concerning the library. In this way, it was impossible to use the content of the statements as a cue in the matching task.

Following S. E. Taylor and colleagues (1978), participants were instructed at the beginning of the experiment to try to form an impression of the group as a whole during the discussion. After the initial instruction phase, participants observed a discussion between the eight male targets who made six statements each, either in the Face condition or in the Voice condition. During the subsequent matching task, 48 new statements were presented alongside the 48 statements from the discussion, and participants first decided whether a given statement was old (i.e., previously presented) or new (i.e., not previously presented). If the statement was categorized as new, then another statement would appear on the screen, whereas if it was categorized as old, then the pictures of all targets from the discussion phase would appear, and participants were asked to indicate which target made this statement. Importantly, this implies that faces were present both in the discussion and in the matching task, whereas voices were only heard during the discussion phase.

## Results

We analyzed the results using the multinomial model validated by Klauer and Wegener (1998). Multinomial models can be applied to any categorical data, and they allow estimating underlying or latent cognitive processes (for an extended introduction, see Klauer & Wegener, 1998, Appendix A). Each cognitive process that is supposed to occur is represented with a certain parameter that is estimated on the basis of frequencies obtained from the data set. Each response is coded for one of the possible types of answers. For example, if a given statement was from an Italian target, it could be assigned either to him correctly, or to another Italian target, or to a German one, or it could be categorized as a new statement. So there are three sources (i.e., Italian target, German target, and new statement), and each statement from each source can be assigned to one of four possible sources (i.e., correct speaker, Italian target, German target, and new statement).[3]

Figure 1 displays one tree of the multinomial processing tree representation (Hu & Batchelder, 1994; Klauer & Wegener, 1998) that describes the participants' answers based on the hypothesized underlying processes of item discrimination (parameter $D$), person discrimination (parameter $c$), and category discrimination (parameter $d$) as well as three guessing processes. The three different guessing processes are as follows: guessing the item status, old versus new (parameter $b$); guessing the category (parameter $a$); guessing the correct person within a given category (parameter $1/n$). The units of measurement for the analyses are the frequencies for each possible assignment (see Tables A1 and A2 for the data matrix and frequencies). For instance, the total number of correct assignments of statements to speakers is hypothesized to result from four different combinations of cognitive processes: either by correctly remembering the statement and speaker, or by remembering the target category and making a lucky guess on the target, or by guessing the target category and the target, or even by guessing that the statement was old and then guessing the target category and the target, as shown in Figure 1. Of course, a multinomial model is only useful to the degree that its parameters are valid representations of the underlying cognitive processes, as this was demonstrated in an extended set of experiments by Klauer and Wegener (1998).

To analyze our data, we used the AppleTree software (Rothkegel, 1999). Our multinomial model consisted of six trees, each corresponding to a different source (Italian speaking, German speaking, new source for the Voice condition; Italian looking, German looking, and new source for the Face condition). The goodness of fit of this model was evaluated by means of chi-square tests. All 19 parameters were estimated using the maximum likelihood method as explained by Batchelder and Riefer (1999) to obtain the baseline model. The baseline model[4] (see Table 1) had very satisfactory goodness of fit, $\chi^2(3, N = 5760) = 1.55, p = .67$, which was the precondition for testing our hypotheses.

We then tested alternative models against the baseline model to see whether there were differences between the parameters. Nonsignificant values of goodness of fit indicate that the deviation of model predictions from the data is not significant (i.e., that the model fits the data), indicating that the respective parameters do not differ significantly. After setting item recognition (parameter $D$) equal for the Face and Voice condition, the model did not fit the

[2] We thank Karl Christoph Klauer for generously sharing his materials with us.

[3] A new statement can either be assigned to a wrong target or classified correctly as a new statement but never to the correct speaker, because it was not made during the discussion. So for new statements, only three assignments are possible.

[4] In order to obtain a baseline model and test whether it fits the data, some parameters had to be set equal. To begin with, we set parameters (corresponding to cognitive processes) for item discrimination equal within each experimental condition (i.e., $D_I = D_G = D_N$). Also, guessing the item status (parameter $b$) was set equal between the two conditions. Finally, the parameter corresponding to guessing the person within the correct category ($1/n$) was kept constant at .25 throughout the categories and conditions.

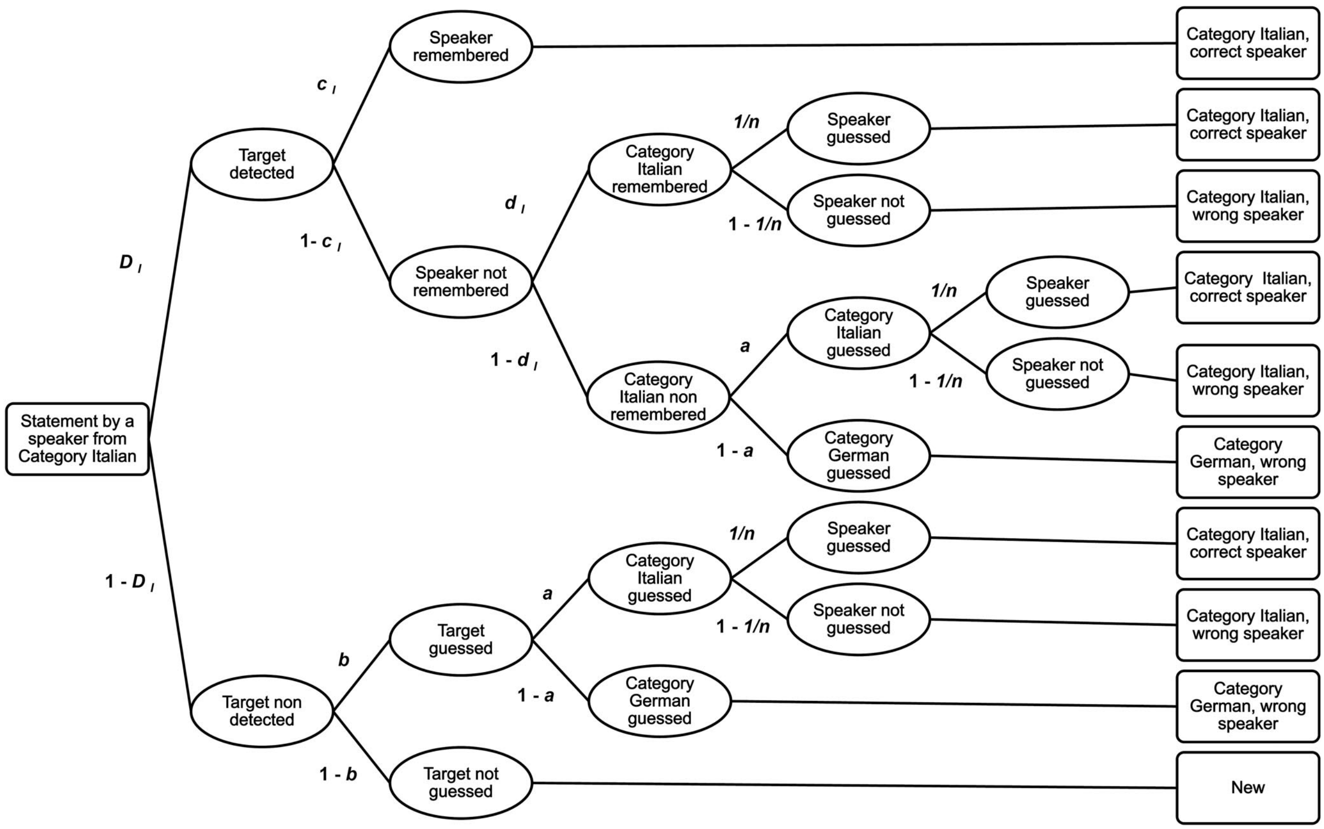


*Figure 1.* Illustration of the multinomial model of social categorization in the modified "Who Said What?" paradigm; partial tree for statements made by an Italian target. Response categories are shown in rectangles to the right. $D_I$ = probability of detecting a statement made by a speaker from Category Italian; $c_I$ = person discrimination parameter; $d_I$ = probability of correctly discriminating category of statement made by a speaker from Category Italian; $a$ = probability of guessing that a statement is made by a speaker from Category Italian; $b$ = probability of guessing that a statement is old.

data anymore, $\Delta\chi^2(1, N = 5760) = 8.4$, $p < .05$, indicating that item memory was superior in the Face than in the Voice condition. Though not predicted, this finding can be explained with the fact that in the Face condition, faces as well as written statements were present both during the discussion as well as in the matching task, facilitating memory retrieval over the Voice condition. Furthermore, setting person memory (parameter $C$) equal between the two conditions, with $\Delta\chi^2(3, N = 5760) = 6.8$, $p = .08$, indicated that person memory was somewhat better for the Face ($C = .30$) than the Voice condition ($C = .23$). There were no other differences in person memory concerning Italian and German targets, $\Delta\chi^2(2, N = 5760) = 2.11$, $p = .55$.

Most important, there was no difference between category memory (given by parameter $d$) in the Face and Voice condition, $\Delta\chi^2(3, N = 5760) = 1.03$, $p = .79$. This indicates that category activation was independent of the stimuli used, that is, whether ethnicity was presented in a visual (i.e., typical Italian- and German-looking men) or auditory (i.e., speaking standard German or with an Italian accent) manner. As is evident from the estimates and confidence intervals shown in Table 1, category memory was clearly different from zero, and the probability of remembering the category was similar for Italian and German targets both in the Face and Voice condition. Hence, we can conclude that the two conditions can be considered equal (i.e., looks and accents activate ethnic categorization to the same degree).

## Discussion

The main finding of Experiment 1 was that ethnicity (German vs. Italian in our case) was equally well presented with visual and auditory stimuli. Given our stimuli, target categorization was not influenced by the fact that ethnicity was presented with looks or with accents. At the same time, better item discrimination in the Face condition compared with the Voice condition indicated that the presence of faces during both the discussion and matching task facilitated the recognition of old statements. In contrast, in the Voice condition, accents were present only during the discussion, resulting in increased difficulty of the matching task. Whereas target categorization based on typical Italian versus typical German looks provides a conceptual replication of previous findings using prototypical looks to indicate ethnicity, ours is the first demonstration that different accents present during a discussion influence social categorization observed at a later point in time. In summary, with the use of the adjusted WSW procedure and using

Table 1
*Parameter Estimation and 90% Confidence Intervals (CIs) for Baseline Model Including Both Conditions (Only Face and Only Voice) in Experiment 1*

| | Face condition | | Voice condition | |
|---|---|---|---|---|
| Parameter | Estimate | CI | Estimate | CI |
| $D_I = D_G$ | .72[a] | .69, .75 | .63[b] | .60, .66 |
| $D_N$ | .72[a] | .69, .75 | .63[b] | .60, .66 |
| $d_I$ | .49 | .33, .66 | .51 | .32, .69 |
| $d_G$ | .48 | .29, .66 | .56 | .42, .69 |
| $c_I$ | .30 | .25, .36 | .21 | .16, .27 |
| $c_G$ | .29 | .24, .35 | .27 | .21, .33 |
| $a$ | .48 | .36, .61 | .57 | .46, .68 |
| $b$ | .16[c] | .14, .19 | .16[c] | .14, .19 |

*Note*. I = Italian targets; G = German targets; N = distractors; $D_I$, $D_G$, $D_N$ = parameters of item discrimination; $d_I$, $d_G$ = parameters of category discrimination; $c_I$, $c_G$ = parameters of person discrimination; $a$ = probability of guessing Category Italian; $b$ = probability of guessing a statement is old. Parameters with the same superscript were set equal.
[a] Parameters set equal. [b] Parameters set equal. [c] Parameters set equal.

a validated multinomial model, Experiment 1 demonstrated that both looks and accent can equally trigger social categorization.

## Experiment 2

The aim of Experiment 2 was to test the influence of looks and accents, when combined, on ethnic categorization. Hence, for some of the targets there was congruent ethnic information (e.g., an Italian-looking male who speaks German with an Italian accent), and for some, it was incongruent (e.g., an Italian-looking male who speaks standard German). Although we are not actually using two different categories, our design can be compared with those looking at cross-categorization.

Data from Experiment 1 suggest that, given our stimuli, both looks and accent are equally salient, and thus participants should tend to categorize on a subgroup level (i.e., taking into account both looks and accent). In contrast, ELIT indicates that language is the most important characteristic of one’s ethnic identity (e.g., Giles, 1977; D. M. Taylor et al., 1973). Therefore, given both looks and accents as indicators of ethnicity, accents should be perceived as more meaningful than looks, and as a consequence should be used to categorize the targets. In other words, according to ELIT, we expect that it should become rather irrelevant to participants what the targets look like. Thus, the activation and encoding on the basis of accents would inhibit the use of information about looks (cf. Macrae & Bodenhausen, 2000). Our prediction is in line with some previous findings (cf. Kurzban et al., 2001), indicating that when a basketball team was made salient, the target race became “invisible.” The structural difference between previous findings and our experiment is that we are not dealing with cross-categorization (along different social categories), but different sources of information on the same category.

Two contrasting hypotheses can be deduced from the cross-categorization literature, on the one hand, and from ELIT, on the other. Following the cross-categorization literature, one would expect the biggest difference between within-accent and within-looks category errors, on the one hand, and on the other, between-accent and between-looks category errors, with the former error type also being the most numerous compared with all possible errors. In other words, participants should tend to find the best fit between individuating and categorizing the target by creating a subgroup (that includes information from both looks and accent). In contrast, based on ELIT, we would predict an interaction between category and error type. To be precise, participants should more often confuse targets belonging to the same accent category independently of their looks category. The strongest empirical evidence for ELIT would be obtained if remembering accents happened as often as remembering both sources of information.

In an additional prestudy ($N = 29$) with the traditional WSW paradigm, we counterbalanced the specific face–voice pairs across conditions (e.g., in one condition, a given Italian-looking face was paired with standard German, and in the other with an Italian accent) in order to ensure that specific face–voice pairs do not influence categorization. Additionally, we counterbalanced which set of statements was used during the discussion part (the “old statements” vs. the “new statements” from Experiment 1). Neither face–voice pair nor set yielded main effects or interactions in any analysis (all $F$s $< 1$). Hence, specific face–voice pairings did not influence our findings. However, we found that participants made significantly more within looks-within accent and between looks-within accent errors than any other error type, consistent with the predictions based on ELIT. Because this method leaves some room for alternative interpretations of findings, we used again a multinomial model in Experiment 2.

In Experiment 2, we tested the underlying processes contributing to online categorization when both looks and accents allow for social categorization. Our main hypothesis was that participants accentuate the accent category and assimilate different looks categories. When comparing category memory based on accents with category memory based on looks, the former should significantly exceed the latter. Additionally, we tested whether category memory based on looks is significantly different from zero.

### Method

**Participants.** Participants were 33 students (21 women and 12 men) from the Friedrich Schiller University Jena with different majors. Their age ranged from 19 to 30 years ($M = 22.64$, $SD = 2.76$). Students took part in the experiment in individual sessions of 20–25 min. At the end of the experiment, they were fully debriefed and rewarded with either a chocolate bar or partial course credit.

**Materials and procedure.** Materials were identical to those used in Experiment 1. There were four types of targets, with two targets for each type (based on different look–accent combinations) involved in a discussion. The procedure was generally identical to Experiment 1; however, 96 statements were used in the matching task (48 new statements and 48 statements known from the discussion part). Because the additional prestudy found no effects of statement set, only one sample of old–new statements was used. The division of phrases as old or new was arbitrary.

### Results

As is shown in Table B1, depending on five different sources (four types of targets and new statements) assignment to six different sources is possible: correct assignment or assignment to any of the five sources. The parameters, or contributing cognitive processes, are similar to those already introduced in Experiment 1 with additions due to the increase in categories.

Specifically, the already known parameters were distractor (*DN*) and item detection ($D_{ii}$, $D_{ig}$, $D_{gi}$, $D_{gg}$),[5] item status guessing (parameter *b*), person memory (parameter *c*), and guessing the person within the relative subgroup (fixed parameter 1/*n*). Additionally, there were several types of category memory in addition to reconstructive category guessing. For *category memory*, we differentiated between three different cognitive processes, represented by three different parameters: Parameter *dvf* (looks, accent: subgroup memory), that is, joined memory for both the looks and accent category; parameter *df* (looks, not accent) exclusive memory for the looks category; and finally, parameter *dvnf* (not looks, but accent), exclusive accent memory. *Reconstructive category guessing*, represented by parameter *a*, corresponds to the bias to choose a speaker from a certain social category instead of another, when there is no actual memory for this specific speaker or his category. The two *a* parameters signify category guessing based on two sources, namely, category guessing based on accents and category guessing based on looks.

Units of analysis were the frequencies of the 29 possible answers (see Table B1 for the data matrix). From the initial model with 38 parameters, we computed the best fitting baseline model that had a satisfactory goodness of fit, $\chi^2(3, N = 3168) = 1.05$, $p = .79$ (see Table 2).[6] Following the rationale of our model, we started by testing person memory (parameter *c*) for the same accent group. After setting parameters equal, there was no significant loss of fit, $\Delta\chi^2(2, N = 3168) = 4.97$, $p = .08$, indicating that there was no statistically significant difference in person memory (i.e., Italian- vs. German-looking targets) within the same accent group. However, after constraining all person memory parameters to be equal, there was a significant loss in fit, $\Delta\chi^2(1, N = 3168) = 27.56$, $p < .001$. This finding indicates that targets who spoke German with an Italian accent ($c = .18$) were remembered less well than targets who spoke standard German ($c = .41$). Even though rather low, target memory for targets with an Italian accent was significantly different from zero, $\Delta\chi^2(1, N = 3168) = 30.93$, $p < .001$. This failure to remember the Italian accent-speaking targets on an individual basis compared with the standard German-speaking targets indicates an outgroup homogeneity effect. Interestingly, this effect was found only based on the way the targets spoke, but independently of their looks. The lack of outgroup homogeneity effect based on looks is a first point of support for our hypothesis that accents inhibit ethnic categorization based on looks.

To actually test the supremacy of accents in ethnic categorization, we compared the different category memory parameters in a stepwise manner. To begin with, we set parameters equal for three different types of category memory. Specifically, after setting equal the parameters for exclusive looks memory for the different target types, $\Delta\chi^2(3, N = 3168) = 1.24$, $p = .74$, for subgroup memory, $\Delta\chi^2(3, N = 3168) = 1.21$, $p = .75$, and finally exclusive accent memory, $\Delta\chi^2(3, N = 3168) = 2.66$, $p = .44$, the model preserved satisfactory goodness of fit. Thus, there was no difference between memories for the different target types. Subsequently, exclusive looks ($df = .15$) and subgroup memory ($dvf = .01$) were set equal, resulting in no significant loss of model fit,

Table 2
*Parameter Estimates and 90% Confidence Intervals (CIs) for Experiment 2*

| Process | Parameter | Estimates | CI |
|---|---|---|---|
| Distractor recognition | $D_N$ | .75[a] | .72, .78 |
| Item recognition | $D_{gg}$ | .75[a] | .72, .78 |
| | $D_{ig}$ | .75[a] | .72, .78 |
| | $D_{gi}$ | .66[b] | .62, .70 |
| | $D_{ii}$ | .66[b] | .62, .70 |
| Item status guessing as old | $b$ | .17[c] | .13, .20 |
| | $b_{gg}$ | .17[c] | .13, .20 |
| | $b_{ig}$ | .17[c] | .13, .20 |
| | $b_{gi}$ | .17[c] | .13, .20 |
| | $b_{ii}$ | .17[c] | .13, .20 |
| Person memory | $c_{gg}$ | .47 | .39, .55 |
| | $c_{gi}$ | .15 | .06, .24 |
| | $c_{ig}$ | .36 | .28, .44 |
| | $c_{ii}$ | .22 | .13, .31 |
| Exclusive accent memory | $d_{gg}vnf$ | .58 | .32, .83 |
| | $d_{gi}vnf$ | .64 | .36, .92 |
| | $d_{ig}vnf$ | .49 | .28, .70 |
| | $d_{ii}vnf$ | .51 | .19, .83 |
| Exclusive looks memory | $d_{gg}f$ | .18 | .01, .36 |
| | $d_{gi}f$ | .09 | .07, .25 |
| | $d_{ig}f$ | .19 | .07, .45 |
| | $d_{ii}f$ | .23 | .02, .49 |
| Subgroup memory | $d_{gg}vf$ | .01 | .002, .02 |
| | $d_{gi}vf$ | .09 | .003, .24 |
| | $d_{ig}vf$ | .11 | .09, .19 |
| | $d_{ii}vf$ | .25 | .06, .46 |
| Category guessing accent | $a_{gg}1$ | .63[d] | .51, .75 |
| | $a_{gi}1$ | .63[d] | .51, .75 |
| | $a_{ig}1$ | .63[d] | .51, .75 |
| | $a_{ii}1$ | .63[d] | .51, .75 |
| | $avf$ | .63[d] | .51, .75 |
| | $avnf$ | .63[d] | .51, .75 |
| Category guessing looks | $af$ | .47[e] | .35, .59 |
| | $a_{gg}2$ | .47[e] | .35, .59 |
| | $a_{gi}2$ | .47[e] | .35, .59 |
| | $a_{ig}2$ | .47[e] | .35, .59 |
| | $a_{ii}2$ | .47[e] | .35, .59 |
| Person guessing | $1/n$ | .50 | constant |

*Note.* $D_N$ = parameter of item discrimination; *gg* = target with German looks who spoke standard German; *ig* = target with Italian looks who spoke standard German; *gi* = target with German looks who spoke German with Italian accent; *ii* = target with Italian looks who spoke German with Italian accent.
[a] Parameters set equal. [b] Parameters set equal. [c] Parameters set equal. [d] Parameters set equal. [e] Parameters set equal.

[5] The abbreviations for four types of item memory are as follows: $D_{ii}$ = Italian-looking and Italian "speaking" target; $D_{ig}$ = Italian-looking and German "speaking" target; $D_{gi}$ = German-looking and Italian "speaking" target; $D_{gg}$ = German-looking and German "speaking" target. This type of abbreviation is kept constant throughout the present article.

[6] In order to gain degrees of freedom, we set some parameters equal. This procedure followed the rationale explained in Klauer et al. (2003) and was adapted to our specific research questions. First, we set equal the parameters for all German-speaking targets to the distractor recognition parameter (i.e., $D_{ig}$, $D_{gg}$, *DN*). Separately, the additional parameters *D* for all Italian-speaking targets were set equal (setting all item-guessing parameters equal did not result in a satisfactory goodness of fit). Additional degrees of freedom were obtained by constraining all parameters for guessing the category looks to be equal (for all subgroups); the same was done for guessing the category accent (for all subgroups). Finally, we set all parameters for item status guessing (parameter *b*) equal.

$\Delta\chi^2(1, N = 3168) = 0.20$, $p = .65$. More important, after setting equal the joint parameter of exclusive looks and subgroup memory with exclusive accent memory, there was a significant loss of goodness of fit, $\Delta\chi^2(1, N = 3168) = 39.38$, $p < .001$, indicating better exclusive memory for accents ($dvnf = .46$) than for looks ($df = .15$), in line with our main hypothesis. Specifically, regarding memory for social categories, participants made no differentiation between different types of targets, but they did discriminate them on the basis of their accents to a significantly stronger degree than on the basis of their looks. Exclusive looks memory was small, but differed significantly from zero, $\Delta\chi^2(1, N = 3168) = 10.70$, $p < .001$, indicating that even though significantly lower than accent memory, there was some face category memory.

## Discussion

The results of Experiment 2 confirmed the hypothesis based on ELIT: Accents provide more relevant information in directing categorization when compared with looks. Accordingly, participants relied mainly on how targets spoke to create categories. This was the case even though the targets' looks were presented in the discussion and matching task, whereas accent was presented only during the discussion part. This shows that the initial information about targets was so important to the participants that it persisted till the end of the matching task, even in the absence of accents during this task. In addition, person memory shows the same pattern, demonstrating that it was practically irrelevant for the participants what the target looked like; instead, they relied only on the way the target spoke to assign him to a given category. This categorization was spontaneous, because participants were only told to observe a discussion and to form an impression of targets as a group. So even if not asked to, participants still categorized the targets on the basis of ethnicity, and they did so using the most important information to them, namely, accent. Unlike person memory, category memory did not differ between different targets (i.e., German or Italian), though category memory for accents was superior to category memory for faces. Additional evidence of the importance of accents in ethnic perception and categorization was found in the outgroup homogeneity effect, within the person memory based on accent (i.e., lower for targets speaking with an Italian accent). The fact that we were able to show these findings using the validated multinomial model rules out the potential concern that they are due to stimulus sampling because that has been shown to affect person memory, not category memory.

# General Discussion

Social categorization seems to be a spontaneous social phenomenon, often underlying other processes such as stereotyping. In the present article, we offered an original approach to this subject by using accents as cues for categorization. The category ethnicity is considered one of the primary social categories next to gender and age. Additionally, it is linked very strongly theoretically to language (e.g., Giles, 1977; D. M. Taylor et al., 1973).

Our first experiment set the stage for the main one. Using a WSW paradigm, we demonstrated that the category ethnicity is spontaneously and equally strongly activated when using visual (i.e., typical Italian vs. German look) or auditory stimuli (i.e., speaking standard German vs. German with an Italian accent). We thus provided empirical evidence for a language-ethnicity bond by showing that social targets were categorized into "standard German speakers" versus "speakers with an Italian accent." To our knowledge, this is the first study demonstrating that accents heard during a discussion provide the basis for social categorization obtained during a matching task in the absence of those accents.

In the second experiment, we investigated the outcome of categorization based on ethnicity by crossing visual (typical looks) and auditory category information (accents). In this way, there was ethnic information both in visual and auditory stimuli, though category information on the two channels was not always congruent (e.g., there were German-looking targets speaking with an Italian accent). On the basis of ethnolinguistic identity theory, ELIT (Giles et al., 1977; Giles & Coupland, 1991; Giles & Johnson, 1981, 1987), we predicted that categorization would be driven by accents (i.e., language). According to our results, it was rather irrelevant for participants what targets looked like; it mainly mattered whether they were speaking with an accent or not. In this case, it was almost as if participants became blind to the visual category information in the presence of more meaningful auditory category information. This pattern of findings was shown using parameter estimates in a multinomial model analysis.

One could argue that the salience of these two presentation modalities may be different to start with. Despite the fact that intuitively, one could say that a typical German and an Italian-looking target differ less than two speakers who speak with or without an accent, the results of Experiment 1 indicated otherwise. When only looks or only accent indicated ethnicity, categorization based on looks appeared equally strong as categorization based on accents. Our pretests also indicated that, subjectively, the difference between the two groups with regard to looks was identical to their difference with regard to accents. Moreover, looks actually had an advantage over accents in our paradigm because they were present in both tasks, during the discussion and during matching, whereas the accent was present only during the discussion. This indicates that targets were categorized online during the discussion, and by the time the matching task came, participants did not need the accent for social categorization anymore. Future studies should test the generalizability of these findings in different contexts, such as with use of African American and White faces in addition to different accents. One could argue that in this case, visual differences would exceed differences between accents. In contrast, based on the finding that children prefer friends who belong to the same accent group independent of their race (Kinzler et al., 2009), one could predict the same pattern of the results as reported above. In a European context, a meaningful comparison would be between African American and White faces, on the one hand, crossed with American versus French accents. We think it is an open question whether the visual differences would indeed be the more meaningful basis for categorization in that case.

A surprising finding of our pretest was that people were hardly able to tell whether two different statements heard in immediate succession were made by the same or different speakers. This points out that people are generally very bad at distinguishing unknown voices from one another. It is often assumed that unknown faces can easily be recognized. Instead, recent research has shown that this is not true (e.g., Hancock, Bruce, & Burton, 2000; Megreya & Burton, 2008). If the same picture is shown twice, the person is easily recognized. However, for two different pictures of the same unfamiliar person, recognition rates were extremely low. Basic research that compares voice and face perception with unknown stimuli still needs to be conducted.

Given our main conclusion that accents appear to mitigate the role of looks for ethnicity categorization, it is crucial to discuss in how far this finding might be limited to the specific stimuli used here. First, it is important to ensure that the two categories of faces used are not closer to each other on the true underlying face dimension than the accents are on the true underlying accent dimension. We have two pieces of evidence that speak against this potential pitfall: On the one hand, our pretest ratings on how German and Italian each face and accent appeared did not differ between faces and accents, demonstrating that subjective ratings of the between-category distances are comparable for faces and accents. On the other hand, in Experiment 1 and two further preexperiments, in the presence of only the faces or only the accents, ethnicity categorization was almost identical along both dimensions (and far from a ceiling or floor effect), yielding experimental evidence that people comparably relied on each dimension for categorization. We presently cannot think of stronger evidence that the distances between categories are similar.

A second potential stimulus-sampling problem is more difficult to rule out: If the within-category similarity is smaller for voices/accents than for faces, category memory could be enhanced for the former dimension because the difficulty to distinguish individuals within each category is increased. This is a serious potential limitation of the traditional WSW paradigm procedure comparing only within- and between-category errors, but it is ruled out in Experiment 2 by using the multinomial model of the WSW paradigm for data analysis. Increasing or decreasing the within-category similarity selectively affects person memory and leaves category memory unchanged (Klauer & Wegener, 1998, Experiment 5). Our interpretations rest precisely on this category memory parameter. In this regard, our research is a perfect demonstration of just how useful model-based analyses are: Unless a model is used where category memory can be estimated separately from person memory, it is very difficult to conclude that one categorical dimension affects memory more than another. An additional illustration for this is our finding that in the pretest rating of voices, German native speakers were perceived as more similar than Italian native speakers. In contrast, our person memory (Experiment 2) findings indicated an outgroup homogeneity effect based on accent: Person memory was lower for those speaking with an Italian accent, demonstrating that not objective stimulus features but sociopsychological processes determine memorability.

What can we conclude from our findings? Clearly, language is a very important and extremely powerful tool in person perception and categorization, and it is diagnostic of many characteristics (e.g., ethnic background, personality traits, etc.). As the Romanian philosopher Emil Cioran (1987) once said: “We inhabit a language rather than a country” (section 7). It becomes apparent, then, that manipulating language can provide new insights into processes of categorization, stereotyping, and even person perception (cf. Schweinberger, Robertson, & Kaufmann, 2007). Schweinberger et al. showed that familiar voices are more easily recognized when presented with time-synchronized articulating faces rather than mere face pictures, indicating that audiovisual integration is complex and of central importance in understanding how we perceive others.

Our findings provide strong support for ELIT (Giles, 1977; Giles & Johnson, 1981, 1987). Over and above providing empirical evidence for the language-ethnicity bond, we were also able to show that in case of categorization, language (i.e., accents) plays the central role. It has long been known that individuals change their manner of speaking to make their group identification more or less salient and that their attitudes and usage of a given language are strongly linked to identification with a group using this language. Here, we went one step farther and showed that language cues are used to make sense of others and to categorize them. Indeed, what better and more reliable way of categorization could there be than to listen to language, an inevitable medium of almost every communication? At this point, it remains unclear whether people who have knowledge of the Italian language would be more precise in categorizing than those who do not. In our case, participants did not have any, or only very little, knowledge and experience with the Italian language. However, chances are that people in possession of more differentiated knowledge would be better at detecting smaller differences between, for example, Italian and Spanish accents in German, but both would be infallible in detecting divergences from standard German in the direction of another language (e.g., Spanish or Italian) and divergences that are made within their native language (e.g., a Bavarian dialect). In other words, future studies should try to test the influence of individual expertise in categorization. It is realistic to assume that people having expertise in a dialect or language are more sensitive in perceiving even smaller variations, whereas for nonexperts, these variations would probably pass unnoticed.

Our work provides a basis for future studies on (cross-) categorization. More important, we show that a given category, in our case ethnicity, can be presented with visual as well as auditory stimuli. Still, accents appear to be the “privileged” channel for indicating ethnicity. Bearing this in mind, it is undoubtedly important to include accents in future studies, because they seem to be very natural, subtle, and yet extremely powerful agents of social perception and categorization.

To date, there has been little research about multichannel presentation of the same category information. To some extent, different category presentations have been used by Klauer and colleagues (2003). They varied the context relevance (e.g., Hewstone et al., 1991; van Knippenberg et al., 1994) of the categories age and gender either by varying stereotypical statements that targets made or by showing pictures clearly indicating different targets (e.g., old-male). When using visual stimuli to present categories, exclusive gender memory was superior to age memory, but the opposite was true when categories were presented verbally (i.e., in the form of stereotypical statements). This finding shows that the stimulus type for category presentation (in this case, context relevance by either visual or verbal stimuli) also influences the outcome of cross-categorization. Similarly, Kurzban and colleagues (2001) showed that race can be erased by making basketball teams salient (with the use of pro arguments concerning a given topic and by sharing the same outfit). We extended their findings by showing that typical looks (German vs. Italian) can be erased quite well if in addition, language cues are provided. In other words, whereas Kurzban et al. (2001) showed that a cross-category can be erased by making a different category salient, we showed that even information pertaining to the same category (ethnicity) can be almost erased if that category is presented on a more meaningful channel (accent).

Our findings are in line with the theory that postulates the use of the first meaningful category (or information) and the inhibition of less meaningful ones (cf. Stangor et al., 1992). Apparently, language is a more solid basis for social categorization than looks, at least in the present case. Admittedly, this might differ given Black versus White targets who speak with typical Black versus White accents. Following our findings, it might be assumed that preexisting cognitive prototypes could be adjusted on the basis of strong, even conflicting information provided by the accent. In other words, the face that was initially perceived as prototypically German might be assimilated with other "Italian" targets due to the prototypical Italian accent. This would lead to more differentiated representations of a given category, Italian in our case (i.e., it would be activated that not all Italians have dark eyes and hair).

Recent research is increasingly taking into account that individuals belong to multiple social categories at the same time (Crisp & Hewstone, 2007). Some of these categories are correlated, whereas others are independent, and all form a hierarchy (Hornsey & Hogg, 2000; Mummendey & Wenzel, 1999)—where those who belong to the same category on one level (e.g., Germans) belong to different categories on a lower level (e.g., Bavarians vs. northern Germans). Thus, the social-constructionist nature of social categories must be taken into account in respective research. By implication, the simplified stimuli presented in empirical investigations may create the very circumstances under which certain social categories become most important. For instance, pictures of fair-haired persons versus pictures of dark-looking persons presented in an experiment may provoke an ingroup–outgroup categorization effect that would not be replicated given the complexity of real social encounters, as our findings demonstrate.

A specificity of accents as a source of information is that accents allow much more differentiation within an ingroup than looks do. Whereas all our "German" speakers spoke standard German, a more realistic scenario would be varying their accents, too. For instance, in a U.S. context, it is unclear whether ingroup–outgroup social categorization based on accents is observed at all if ingroup members speak with a midwestern, southern, or eastern accent. In other words, it is possible that the social categories that appear to be privileged are not privileged due to their importance for people, but due to the materials used in categorization studies. The principal advantage that accents offer is their subtle, yet very differentiated presence in everyday communication. Therefore, accents give researchers the possibility to make more distinctive groups without making them appear artificial, because hearing different regional accents, including, for example, a Bavarian accent in German, is more common than seeing people wearing leather trousers while drinking huge beers (e.g., in the case of stereotypical Bavarians). Future research using accents will show whether established findings generalize to these stimuli and in how far theories on social categorization need to be refined to incorporate these findings.

One reason why visual stimuli have been used in the majority of studies is probably the dominance of the visual sense in humans, mirrored in the dominance of vision research in the field of perception. This is probably why many instruments (e.g., medical, etc.) have visual indicators of measures. Whereas a self-imposed restriction to visual stimuli may make sense in many areas of social psychology, our findings show that social categorization may be dominated by other stimuli, in the present case, auditory ones. Hence, if we are to gain greater external validity to our findings, we should pay attention to the question of what kind of stimuli we use to present certain categories.

What do these findings mean for everyday life? Many migrants try hard to adapt their clothing, behavior, and look to the host country. But even though one may try hard to learn the language of the host country well, an accent from one's native language is extremely hard to lose (v. Humboldt, 1836; see also Gluszek & Dovidio, 2010). Our findings confirm from a social psychological perspective that language is the main avenue to integration, an idea that is well established from a sociological point of view (e.g., Esser, 2006). Even though our targets spoke error-free German, the mere presence of an Italian accent was enough to categorize them as Italians (and therefore, as possible outgroup members). Thus, the threshold for true integration may be extremely high and, according to findings on language learning, not obtainable for adults. To conclude: We should not forget that language is the primary tool we use while regularly communicating with the world outside, and language should therefore be represented more in social psychological studies. Luckily, technical progress today allows improving ecological validity in this way without sacrificing experimental control.

## References

Aboud, F. E., Clement, R., & Taylor, D. M. (1974). Evaluational reactions to discrepancies between social class and language. *Sociometry, 37,* 239–250. doi:10.2307/2786378

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., . . . Schaller, M. (2006). They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity. *Psychological Science, 17,* 836–840. doi:10.1111/j.1467-9280.2006.01790.x

Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Addison-Wesley.

Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *Journal of Social Psychology, 5,* 37–55. doi:10.1080/00224545.1934.9921582

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Bianchi, M., Mummendey, A., Steffens, M. C., & Yzerbyt, V. Y. (2010). What do you mean by "European"? Evidence of spontaneous ingroup projection. *Personality and Social Psychology Bulletin, 36,* 960–974.

Bourhis, R. Y., Giles, H., & Tajfel, H. (1973). Language as a determinant of Welsh identity. *European Journal of Social Psychology, 3,* 447–460. doi:10.1002/ejsp.2420030407

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer Jr. (Eds.), *A dual process model of impression formation* (pp. 1–36). Hillsdale, NJ: Erlbaum.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152. doi:10.1037/h0043805

Cioran, E. M. (1987). Anathemas and admirations. Retrieved from http://en.wikiquote.org/wiki/Emil_Cioran

Crisp, R. J., & Hewstone, M. (2006). Multiple social categorization: Context, process, and social consequences. In R. J. Crisp & M. Hewstone (Eds.), *Multiple social categorization: Processes, models and applications* (pp. 3–22). New York, NY: Psychology Press.

Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 39, pp. 163–254). San Diego, CA: Elsevier Academic Press.

Elwell, C. M., Brown, R. J., & Rutter, D. R. (1984). Effects of accent and visual information on impression formation. *Journal of Language and Social Psychology, 3,* 297–299. doi:10.1177/0261927X8400300404

Esser, H. (2006). Sprache und Integration. Die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten [Language and integration. The social conditions and consequences of language acquisition by migrants]. Frankfurt, Germany: Campus Verlag.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, form category based to individuating processes: Influence of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). San Diego, CA: Academic Press.

Frable, D. E., & Bem, S. L. (1985). If you are gender schematic, all members of the opposite sex look alike. *Journal of Personality and Social Psychology, 49,* 459–468. doi:10.1037/0022-3514.49.2.459

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509–517. doi:10.1037/0022-3514.60.4.509

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics, 15,* 87–105.

Giles, H. (Ed.). (1977). *Language, ethnicity and intergroup relations*. London, England: Academic Press.

Giles, H., & Billings, A. C. (2004). Language attitudes. In A. P. Davies & E. Elder (Eds.), *The handbook of applied linguistics* (pp. 187–209). Oxford, England: Blackwell. doi:10.1002/9780470757000.ch7

Giles, H., Bourhis, R., & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307–348). London, England: Academic Press.

Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences.* Belmont, CA: Thomson Brooks/Cole Publishing.

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,* 69–100. doi:10.1515/ijsl.1987.68.69

Giles, H., & Niedzielski, N. (1998). German sounds awful, but Italian is beautiful. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 85–93). Harmondsworth, England: Penguin.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation.* London, England: Academic Press.

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review, 14,* 214–237. doi:10.1177/1088868309359288

Gudykunst, W. B., & Ting-Toomey, S. (1990). Ethnic identity, language and communication breakdowns. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 309–327). Chichester, England: Wiley.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4,* 330–337. doi:10.1016/S1364-6613(00)01519-9

Hewstone, M., Hantzi, A., & Johnston, L. (1991). Social categorization and person memory: The pervasiveness of race as an organizing principle. *European Journal of Social Psychology, 21,* 517–528. doi:10.1002/ejsp.2420210606

Hornsey, M. J., & Hogg, M. A. (2000). Assimilation and diversity: An integrative model of subgroup relations. *Personality and Social Psychology Review, 4,* 143–156. doi:10.1207/S15327957PSPR0402_03

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of engineering processing tree models with the EM algorithm. *Psychometrika, 59,* 21–47. doi:10.1007/BF02294263

Kamachi, M., Hill, H., Lander, K., & Vatikiotis-Bateson, E. (2003). "Putting the face to the voice": Matching identity across modality. *Current Biology, 13,* 1709–1714. doi:10.1016/j.cub.2003.09.005

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 104,* 12577–12580.

Kinzler, K. D., Shutts, K., Dejesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27,* 623–634. doi:10.1521/soco.2009.27.4.623

Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effects patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology, 39,* 332–354. doi:10.1016/S0022-1031(03)00017-9

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75,* 1155–1178. doi:10.1037/0022-3514.75.5.1155

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the voice reveals: Within- and between-category stereotyping on the basis of voice. *Personality and Social Psychology Bulletin, 32,* 806–819. doi:10.1177/0146167206286627

Ko, S. J., Judd, C. M., & Stapel, D. A. (2009). Stereotyping based on voice in the presence of individuating information: Vocal femininity affects perceived competence but not warmth. *Personality and Social Psychology Bulletin, 35,* 198–211. doi:10.1177/0146167208326477

Krueger, J. (1996). Probabilistic national stereotypes. *European Journal of Social Psychology, 26,* 961–980. doi:10.1002/(SICI)1099-0992(199611)26:6<961::AID-EJSP799>3.0.CO;2-F

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98,* 15387–15392. doi:10.1073/pnas.251541498

Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social Issues, 23,* 91–109. doi:10.1111/j.1540-4560.1967.tb00578.x

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51,* 93–120. doi:10.1146/annurev.psych.51.1.93

McNamara, T. F. (1987). Languge and social idenity: Israelies abroad. *Journal of Language and Social Psychology, 6,* 215–228. doi:10.1177/0261927X8763005

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14,* 364–372. doi:10.1037/a0013464

Mummendey, A., Otten, S., Berger, U., & Kessler, T. (2000). Positive-negative asymmetry in social discrimination: Valence of evaluation and salience of categorization. *Personality and Social Psychology Bulletin, 26,* 1258–1270. doi:10.1177/0146167200262007

Mummendey, A., & Wenzel, M. (1999). Social discrimination and tolerance in intergroup relations: Reactions to intergroup difference. *Personality and Social Psychology Review, 3,* 158–174. doi:10.1207/s15327957pspr0302_4

Oakes, P. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. Oakes, S. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 117–141). Oxford, England: Blackwell.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339. doi:10.1037/0033-295X.95.3.318

Rothkegel, R. (1999). AppleTree: A multinomial processing tree modeling program for Macintosh computers. *Behavior Research Methods, Instruments and Computers, 31,* 696–700.

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,* 267–273.

Schweinberger, S. R., Robertson, D., & Kaufmann, J. M. (2007). Hearing facial identities. *The Quarterly Journal of Experimental Psychology, 60,* 1446–1456. doi:10.1080/17470210601063589

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* (pp. 33–56). New York, NY: Wiley.

Sherman, S. J., Castelli, L., & Hamilton, D. L. (2002). The spontaneous use of a group typology as an organizing principle in memory. *Journal of Personality and Social Psychology, 82,* 328–342. doi:10.1037/0022-3514.82.3.328

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207–218. doi:10.1037/0022-3514.62.2.207

Taylor, D. M., Bassili, J. N., & Aboud, F. E. (1973). Dimensions of ethnic identity: An example from Quebec. *Journal of Social Psychology, 89,* 185–192. doi:10.1080/00224545.1973.9922590

Taylor, S. E., & Falcone, H.-T. (1982). Cognitive bases of stereotyping: The relationship between categorization and prejudice. *Personality and Social Psychology Bulletin, 8,* 426–432. doi:10.1177/0146167282083006

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778–793. doi:10.1037/0022-3514.36.7.778

v. Humboldt, W. (1836). Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts [The heterogeneity of language and its influence on the intellectual development of mankind]. In *Ueber die Kawi-Sprache auf der Insel Java* [On the Kawi-Language on Java Island] (pp. I-CCCCXXX). Berlin, Germany: Koeniglichen Akademie der Wissenschaften.

van Knippenberg, A., van Twuyver, M., & Pepels, J. (1994). Factors affecting social categorization processes in memory. *British Journal of Social Psychology, 33,* 419–431.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115–1125.

Wenzel, M., Mummendey, A., & Waldzus, S. (2007). Superordinate identities and intergroup conflict: The ingroup projection model. *European Review of Social Psychology, 18,* 331–372.

Yamauchi, T., & Yu, N.-Y. (2008). Category labels versus feature labels: Category labels polarize inferential predictions. *Memory & Cognition, 36,* 544–553.

Yzerbyt, V. Y., Leyens, J.-P., & Bellour, F. (1995). The ingroup overexclusion effect: Identity concerns in decisions about group membership. *European Journal of Social Psychology, 25,* 1–16.

Yzerbyt, V., Provost, V., & Corneille, O. (2005). Not competent but warm . . . really? Compensatory stereotypes in the French-speaking world. *Group Processes and Intergroup Relations, 8,* 291–308.

Zàrate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition, 8,* 161–185.

Zuckerman, M., & Driver, R. E. (1989). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior, 13,* 67–82.

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.

(*Appendices follow*)

## Appendix A

Table A1
*Source/Assignment Frequencies for the Face Condition in Experiment 1*

| | Assignment to: | | | |
|---|---|---|---|---|
| | | An incorrect target from set of | | |
| Source | The correct target | Italian looks | German looks | The set of new items |
| Italian looks | 1 ($N$ = 229) | 2 ($N$ = 214) | 3 ($N$ = 111) | 4 ($N$ = 166) |
| German looks | 5 ($N$ = 228) | 6 ($N$ = 108) | 7 ($N$ = 217) | 8 ($N$ = 167) |
| New | | 9 ($N$ = 32) | 10 ($N$ = 34) | 11 ($N$ = 1,374) |

*Note*. The response categories are numbered from 1 to 9 with respective frequencies of assignments summed over participants.

Table A2
*Source/Assignment Frequencies for the Voice Condition in Experiment 1*

| | Assignment to: | | | |
|---|---|---|---|---|
| | | An incorrect speaker from set of | | |
| Source | The correct speaker | Italian accent | German accent | The set of new items |
| Italian accent | 1 ($N$ = 186) | 2 ($N$ = 247) | 3 ($N$ = 99) | 4 ($N$ = 188) |
| German accent | 5 ($N$ = 196) | 6 ($N$ = 109) | 7 ($N$ = 206) | 8 ($N$ = 209) |
| New | | 9 ($N$ = 44) | 10 ($N$ = 33) | 11 ($N$ = 1,363) |

*Note*. The response categories are numbered from 1 to 9 with respective frequencies.

## Appendix B

Table B1
*Source/Assignment Frequencies for Experiment 2*

| | Assignment to: | | | | | |
|---|---|---|---|---|---|---|
| | | An incorrect speaker from set of | | | | |
| Source | The correct speaker | Italian looks, Italian accent | German looks, Italian accent | Italian looks, German accent | German looks, German accent | The set of new items |
| Italian looks, Italian accent | 1 ($N$ = 106) | 2 ($N$ = 50) | 3 ($N$ = 74) | 4 ($N$ = 30) | 5 ($N$ = 19) | 6 ($N$ = 118) |
| German looks, Italian accent | 7 ($N$ = 95) | 8 ($N$ = 98) | 9 ($N$ = 54) | 10 ($N$ = 17) | 11 ($N$ = 26) | 12 ($N$ = 106) |
| Italian looks, German accent | 13 ($N$ = 146) | 14 ($N$ = 41) | 15 ($N$ = 32) | 16 ($N$ = 39) | 17 ($N$ = 59) | 18 ($N$ = 80) |
| German looks, German accent | 19 ($N$ = 172) | 20 ($N$ = 21) | 21 ($N$ = 34) | 22 ($N$ = 52) | 23 ($N$ = 33) | 24 ($N$ = 84) |
| New | | 25 ($N$ = 20) | 26 ($N$ = 21) | 27 ($N$ = 11) | 28 ($N$ = 13) | 29 ($N$ = 1,517) |

*Note*. The response categories are numbered from 1 to 29 with respective frequencies.

Received January 19, 2010
Revision received August 12, 2010
Accepted August 14, 2010 ■