A Perpetrator’s Accent Impairs Witnesses’ Memory for Physical Appearance

Kerri L. Pickel · Joshua B. Staller

© American Psychology-Law Society/Division 41 of the American Psychological Association 2011

Abstract Sometimes witnesses to crimes must remember both a perpetrator’s appearance and voice. Drawing upon multiple resource theory as well as previous findings that processing foreign-accented speech is more demanding than processing unaccented speech, we hypothesized that a perpetrator’s accent can impair memory for his or her appearance. In Experiment 1, we used a secondary visual search task to demonstrate that processing an accented versus unaccented message demands more cognitive resources. In two additional experiments, we extended that result by showing that witnesses trying to encode information spoken by a perpetrator with an accent rather than no accent provided poorer physical descriptions of him and identified his voice less accurately. We also found that witnesses who heard a more versus less detailed message reported fewer correct details about the perpetrator’s appearance (Experiment 2), and a more rather than less threatening message led to less accurate descriptions (Experiment 3).

Keywords Eyewitness memory · Eyewitness identification · Earwitness memory · Voice identification · Perpetrator’s accent · Attentional resources

While committing a crime, perpetrators sometimes give commands or instructions to victims or bystander witnesses, or they may converse with an accomplice. In such cases, investigators might ask witnesses to remember not only the perpetrator’s physical appearance but also information related to his or her verbal communication. Witnesses can try to describe the qualities of the perpetrator’s voice and report the content of his or her message, and eventually they might attempt to identify the voice in a lineup.

When witnesses try to remember both visual and auditory information about a perpetrator, it is possible that a variable associated with one modality will influence memory for information presented in the other. For example, a weapon’s presence (a visual variable) can impair memory for the content of the perpetrator’s remarks (an auditory memory; Pickel, French, & Betts, 2003). The purpose of the current research is to search for cross-modal interference in the opposite direction. Specifically, we investigated whether a perpetrator’s accent can impair witnesses’ memory for his or her appearance. This issue has real-world implications; through travel or immigration, for example, witnesses may encounter a perpetrator who speaks with what they perceive as a foreign accent (defined as non-pathological speech that diverges noticeably from native speaker pronunciation norms and includes elements such as phoneme substitutions and distortions and prosodic deviations; Munro & Derwing, 1995b).

Multiple resource theory (Wickens, 2002; 2008) provides a framework for understanding why cross-modal interference could affect eyewitness performance. This model accounts for variability in dual-task performance by proposing that cognitive resources are both limited and allocatable and that they are organized into four distinct dimensions, each with two levels. For example, one dimension involves perceptual modalities, including a visual and an auditory level. If two tasks draw resources from different levels along a particular dimension, time sharing (i.e., performing both tasks concurrently) can often be relatively successful. The two tasks will still compete for common perceptual resources, however, so difficult tasks that demand

K. L. Pickel (✉) · J. B. Staller
Department of Psychological Science, Ball State University, Muncie, IN 47306, USA
e-mail: kpickel@bsu.edu

Published online: 21 January 2011
a large quantity of resources could deplete the available resources within their respective modality-specific pools and require supplementation from the larger common pool. In this way, cross-modal interference could occur as the two tasks compete for common resources. The multiple resource model can be applied to Pickel et al.’s (2003) data as follows. Because a weapon captures attention (i.e., the weapon focus effect; Pickel, Ross, & Truelove, 2006), watching an armed perpetrator with the intent to remember his or her physical appearance is a demanding task. Witnesses could simultaneously listen to and remember the semantic content of an armed perpetrator’s speech only when that content was relatively simple. When it was complex, the visual and auditory tasks competed for supplementary common resources, and witnesses could not complete both tasks successfully. Because the visual task was defined as primary within this study, it was the auditory task that suffered, and witnesses showed noticeable deficits in their ability to remember the content of the complex statement.

Thus, visual stimuli can interfere with witnesses’ memory for auditory information. The reverse should also be true if witnesses define the auditory task as primary, as they would when a perpetrator gives orders or instructions that witnesses must attend to, comprehend, and remember to avoid injury to themselves or others. The auditory task would also have to be demanding for interference to occur. One variable that might cause witnesses to perceive the task as more demanding is the perpetrator’s accent.

Previous research using various performance measures has indicated that listeners must exert more effort to process foreign-accented versus unaccented speech. For example, intelligibility is the degree to which listeners correctly understand an utterance (Munro & Derwing, 1995b) and can be measured by asking listeners to transcribe a target speaker’s words verbatim. Native English-speaking subjects make fewer transcription errors when listening to a native English speaker rather than a native Korean (Bent & Bradlow, 2003) or native Chinese speaker (Bent & Bradlow, 2003; Munro & Derwing, 1995a, 1995b). Another measure is comprehensibility, or listeners’ ratings of the difficulty in understanding the speech. Non-native speakers are often rated as less comprehensible than native speakers (Munro & Derwing, 1995b, 1998). Processing speed can also reflect cognitive effort; Munro and Derwing (1995b) found that native English listeners needed more time to verify sentences spoken with a Mandarin accent as opposed to no accent. The authors suggested that a foreign accent imposes “costs” on listeners because they may either fail to recognize some of the speaker’s phonetic segments or words or have to expend extra effort to identify segments or words, for example by recruiting top-down processes or by “replaying” some or all of the message from short-term memory.

If processing accented speech is especially demanding, witnesses should find it harder to identify the voices of targets who speak with an accent (or who are speaking an unfamiliar language, for that matter) as opposed to those with no accent. Because listeners interact frequently with residents of their own geographic region and develop considerable experience with the phoneme pronunciation and prosodic patterns prevalent in their area, they may become expert at detecting subtle variations within the regional “speech schema,” which in turn can help them match a candidate voice in a lineup to a voice held in memory (Goggin, Thompson, Strube, & Simental, 1991). In contrast, accented voices may diverge from the familiar schema so much that listeners find it hard to recognize and utilize small variations within the speech to distinguish one voice from another (and voices speaking a foreign language might be particularly challenging due to even greater departure from the schema). Moreover, the increased costs of processing accented speech may reduce the cognitive resources available for detecting potentially useful cues. Several researchers have reported results consistent with this reasoning. English-speaking listeners can more accurately identify voices speaking unaccented English compared to voices speaking English with a Spanish, German, or French accent (Doty, 1998; Goggin et al., 1991; Thompson, 1987). Furthermore, monolingual listeners can better identify voices speaking their own language than voices speaking another language (Goggin et al., 1991; Philippon, Cherryman, Bull, & Vrij, 2007).

Based on the multiple resource model and previous research on listeners’ performance, we propose that processing accented speech is so demanding that it requires supplemental resources from the common pool along with modality-specific resources. Thus, this auditory task could interfere with a concurrent visual task that is also demanding enough to draw from the common pool. The goal of the present research was to test the hypothesis that witnesses would remember a perpetrator’s appearance less accurately if he or she spoke the witnesses’ native language with an accent rather than no accent. Although the perpetrators in our study were not holding attention-capturing weapons as Pickel et al.’s (2003) were, we thought the task of trying to encode and remember the details of a perpetrator’s appearance would nevertheless be fairly effortful and that performance on this task would reveal cross-modal interference if the perpetrator spoke with an accent versus no accent.

In Experiment 1, we attempted to collect direct evidence that processing accented speech does, in fact, require more resources than processing unaccented speech by instructing participants to perform a secondary task involving visual search while listening to a spoken message. In Experiments 2 and 3, participants acted as witnesses to a carjacking and
a bank robbery, respectively, and subsequently tried to remember the perpetrator’s appearance and message and to identify him in a photo and a voice lineup. In the second and third experiments the perpetrator’s statement contained instructions that the witnesses needed to attend to and remember. Thus, as in a comparable real-life situation, the witnesses made the auditory task their top priority.

**Experiment 1: Does Comprehending a Message Spoken With an Accent Versus No Accent Require More Cognitive Resources?**

The objective of this experiment was to demonstrate that it is especially cognitively demanding for witnesses to listen to a message spoken with an accent with the intent to understand and remember its content. To this end, we asked participants to listen to an audio recording simulating a crime and to prioritize processing of the message spoken by the perpetrator. We reasoned that the amount of resources required to complete this primary task would be reflected in participants’ performance on a concurrent secondary task involving visual search. Specifically, participants from the Midwestern U.S. who hear a statement spoken with an Irish (i.e., foreign) accent should allocate more resources to process it and should therefore perform worse on the secondary task compared to those who hear an unaccented statement.

We were interested in examining the influence of the speaker’s accent using messages with different levels of detail. We expected a message containing a higher level of detail to demand more resources than a less detailed message and therefore to be more difficult to comprehend and remember. Therefore, we anticipated a main effect such that participants in the high detail condition would perform worse on the secondary task. However, we made no specific prediction about whether we would obtain an interaction. We assumed that processing an unaccented message is relatively easy, so participants should usually be able to perform this task simultaneously with visual search. If the amount of detail within the message rises above a critical threshold, though, performance on the secondary task might decline. Thus, we expected a simple effect of detail within the unaccented condition. It is less clear what should happen when processing an accented message, however. The same simple effect could occur, or alternatively the auditory task could be so demanding that changing the amount of detail makes no difference.

Besides asking participants to complete a secondary task, we also required them to remember the perpetrator’s statement. The multiple resource model makes no clear prediction regarding whether accent or amount of detail should affect memory for message content. Because comprehending and remembering the message is primary, poor performance in the accented and/or high detail conditions would merely indicate that the task is too demanding given the available resources and would not reflect cross-modal interference. In other words, participants’ memory for the statement is simply an indicator of their ability to perform the primary task under the different accent and detail conditions.

**Method**

**Participants.** Introductory psychology students (N = 179) at a Midwestern U.S. university participated to fulfill a course requirement. They ranged in age from 18 to 24 years (M = 18.84, SD = 1.06); 71% were female and 90% self-identified as White. All were from the Midwestern United States, spoke English as their first language, and did not speak Irish Gaelic or any related language.

**Materials and Procedure.** We created four versions of a video depicting a carjacking. In all versions, the camera is positioned in the driver’s seat of a car and pointed toward the front passenger’s seat. As the video begins, the passenger door opens and a White male carjacker gets in, looks toward the camera, and gives instructions for driving to an ATM, withdrawing cash, and handing it over (see Appendix A for the script). Although the carjacker claims to have a gun, no weapon is visible.

We manipulated two independent variables. First, the actor portraying the carjacker spoke either with an Irish accent or a Midwestern U.S. accent. To participants from the Midwest, the carjacker would be perceived in the latter condition as not having an accent.

We also manipulated the amount of detail in the carjacker’s message. The Low Message Detail statement included 28 details (the running time of this video was 35 s). In the High Message Detail condition, there were 44 details (running time 55 s). As an example, “in my pocket” (the location of the gun, according to the carjacker) was one detail. The number of details in each version was counted and agreed upon by the authors.

The soundtrack of each video was transferred to an audio CD. Participants were tested individually and were assigned randomly to listen to one of the four audio recordings. There were 43 participants in the Midwest/Low Message Detail condition, 45 in the Midwest/High Message Detail condition, 46 in the Irish/Low Message Detail condition, and 45 in the Irish/High Message Detail condition. They were instructed that their primary goal was to remember the carjacker’s remarks as accurately as possible. While listening to the carjacker speaking, participants completed a secondary task that required them to search for two different targets (a black 8 and a red 2) on a page containing a matrix of randomly ordered digits, with each
digit printed in either black or red ink. Specifically, participants used a pencil to draw a line though each target they could find on the page.

Next, participants completed a written form with two sections. The first section asked for free recall of the target’s message. Participants were asked to provide as much detail as possible. The second section requested demographic information.

The study’s design was 2 (Accent: Irish, Midwestern) × 2 (Amount of message detail: Low, High) factorial. Both variables were manipulated between subjects.

Results

Secondary Task Performance. Because the high and low detail audio recordings differed in length, we measured participants’ performance on the secondary task by dividing the number of targets crossed out by the running time of the recording. Thus, for each participant we calculated the number of targets found per second. A 2 × 2 analysis of variance revealed a main effect of accent, \( F(1, 175) = 6.04, p = .02, \eta^2_p = .03, 95\% CI [.007, .06] \); participants who heard the Midwestern accent performed better \( (M = .32, SD = .10) \) than those who heard the Irish accent \( (M = .29, SD = .11) \). In addition, participants in the Low Message Detail condition found more targets \( (M = .35, SD = .11) \) than participants in the High Message Detail condition \( (M = .26, SD = .09) \), \( F(1, 175) = 34.15, p < .001, \eta^2_p = .16, 95\% CI [.06, .11] \). There was no interaction \( (p = .67) \).

We also examined the number of errors on the secondary task. Most participants did not identify any distracters as targets (overall \( M = .20, SD = .55 \)), and there were no significant effects \( (ps \geq .55) \).

Memory for the Carjacker’s Message. Two independent coders blind to the experimental conditions calculated participants’ memory for the message using the scoring key created by the authors that identified the correct details. Any detail reported by a participant that was not identified on the scoring key was counted as incorrect. The coders tallied the number of correct and incorrect details for each participant. Interrater reliability was high; using a sample of 50 questionnaires, \( r = .97 \).

Before beginning the data analyses, we arbitrarily designated one coder as the “primary coder.” In the case of disagreements, we entered the score figured by the primary coder into the analyses. We felt that adopting this method (which has been used in previous research on eyewitness description accuracy; e.g., Pickel et al., 2003, 2006), rather than arbitrating disagreements ourselves or asking the coders to resolve their disagreements through discussion, would reduce the chance of bias.

The number of details varied across conditions, we analyzed the proportion of correct details provided by each participant (i.e., the number of correct details reported divided by the maximum number that could have been reported as indicated on the scoring key). We obtained a main effect of accent such that participants in the Midwestern condition provided a higher proportion of correct details \( (M = .36, SD = .12) \) compared to participants in the Irish condition \( (M = .31, SD = .13) \), \( F(1, 175) = 7.69, p = .006, \eta^2_p = .04, 95\% CI [.01, .08] \). Moreover, participants who heard the low detail statement reported a higher proportion of correct details \( (M = .40, SD = .12) \) than those who heard the high detail statement \( (M = .27, SD = .10) \), \( F(1, 175) = 62.29, p < .001, \eta^2_p = .26, 95\% CI [.09, .16] \). We obtained no interaction \( (p = .35) \).

Regarding the number of incorrect details, participants who heard the Midwestern accent performed better \( (M = 1.22, SD = 1.07) \) than those who heard the Irish accent \( (M = 1.60, SD = 1.51) \), \( F(1, 175) = 5.17, p = .02, \eta^2_p = .03, 95\% CI [-.76, -.05] \). Additionally, participants in the Low Message Detail condition reported fewer incorrect details \( (M = .88, SD = .89) \) than those in the High Detail condition \( (M = 1.94, SD = 1.46) \), \( F(1, 175) = 36.13, p < .001, \eta^2_p = .17, 95\% CI [-1.42, -.72] \). There was no interaction \( (p = .18) \).

Discussion

Compared to participants in the Midwestern condition, those who heard the accented statement performed worse on the secondary task. These results support the hypothesis that processing accented speech requires more cognitive resources than processing speech that the listener perceives as unaccented. The secondary task results also indicate that participants needed more resources to process the more detailed message compared to the less detailed one. However, we did not find an interaction; increasing the level of message detail from low to high impaired participants’ performance regardless of whether the accent was present.

In addition, participants remembered the message more poorly when it was more rather then less detailed and when the perpetrator spoke with an accent as opposed to no accent. These findings imply that processing the message in these conditions was so demanding that performance declined even though participants allocated supplemental resources to this task at the expense of the secondary task.

Experiment 2: Can Processing an Accented Message Impair Memory for the Speaker’s Appearance?

The results of Experiment 1 provide direct evidence that attempting to comprehend an accented message requires...
more cognitive resources than attempting to comprehend an unaccented message. Also, a message containing a greater level of detail seems to demand more resources. The purpose of Experiment 2 was to go a step further and test the hypothesis that, having allocated additional resources to the task of processing an accented or a high detail message, witnesses’ ability to encode and remember the appearance of the speaker would be reduced compared to conditions with an unaccented or a less detailed message. In this experiment participants watched one version of the carjacking video and subsequently tried to remember the perpetrator’s appearance by providing a physical description and by attempting to identify him in a photo lineup. Based on the first experiment’s data, we predicted main effects of both accent and amount of detail.

We also asked participants to identify the carjacker’s voice. Previous research (Doty, 1998; Goggin et al., 1991; Thompson, 1987) suggests that participants listening to the Irish rather than the Midwestern accent should perform more poorly. We know of no previous research investigating whether the amount of detail in a speaker’s message affects voice identification. However, we reasoned that there would probably be no main effect of this variable, because increasing the level of detail would not change the degree to which the perpetrator’s speech conforms to listeners’ schemas and thus would not harm their ability to judge whether a lineup voice matches the voice held in memory.

Method

Participants. Introductory psychology students (N = 191) at the same university as the subjects in Experiment 1 participated to fulfill a course requirement. They ranged in age from 18 to 58 years (M = 19.59, SD = 3.51), and most were female (63%) and White (84%). All were from the Midwestern U.S., were native English speakers, and did not speak Irish Gaelic or any related language.

Stimulus and Pilot Study. We used the four versions of the carjacking video created for Experiment 1. Previous research has demonstrated that a speaker’s accent can activate a negative stereotype (Munro & Derwing, 1995b; Yarmey, 2007), which in turn could conceivably affect what observers notice and recall about the speaker. Therefore, to ensure that attitudes toward the carjacker were not different between the accent conditions, we conducted a pilot study. Sixty-eight pilot participants, tested in groups of up to 10 individuals, were randomly assigned to watch one of the four videos. Afterwards, they completed a social distance scale (Winer, Bonner, Blaney, & Murray, 1981) that measured their attitudes toward the carjacker. Scores on this scale can range from 6 to 42, with higher numbers reflecting more negative attitudes.

A factorial analysis of variance revealed no main effect of accent; attitudes were similar in the Midwestern (M = 38.76, SD = 4.21) and Irish (M = 37.21, SD = 5.94) conditions, F(1, 64) = 1.70, p = .20, ηp² = .03, 95% CI [-.87, 4.15]. In addition, there was no main effect of amount of message detail and no interaction (p ≥ .23). Not surprisingly, participants’ attitudes were fairly negative across conditions (overall M = 37.99, SD = 5.17).

Procedure. The participants in the main experiment assumed the role of victim-witnesses and watched one of the four versions of the carjacking video. They participated in groups of up to 10 students. Before watching, the experimenter asked them to imagine having stopped their car at a red light when a carjacker suddenly jumps into the vehicle and begins giving orders. As part of the role playing, the witnesses were urged to pay close attention to the carjacker’s words because they would need to be able to follow his instructions to avoid getting hurt.

After watching the video, the witnesses completed a written questionnaire with three sections. First, the witnesses tried to recall the carjacker’s message in as much detail as possible. Next, they attempted to remember his appearance, including both his physical characteristics (e.g., race and build) and his clothing (e.g., shirt and hat). Some items in this section were multiple choice, and some were open-ended (e.g., How would you describe the style of his shirt?). The questionnaire concluded with a demographics section.

Next, the witnesses individually completed two lineup tasks. First, they tried to identify the carjacker in a six-person, target-present, simultaneous photo lineup presented as a 3 × 3 matrix. In constructing the lineup, we followed guidelines established by previous researchers (see Malpass, Tredoux, & McQuiston-Surett, 2007) by (a) selecting fillers who generally resembled the perpetrator but were not so highly similar as to be “clones”; (b) making sure no lineup member stood out from the others on the basis of one or more attributes; (c) including five fillers in addition to the perpetrator; and (d) randomizing the order of the photos for each witness. Witnesses were told the carjacker might not be in the lineup and were given the option of responding that his photo was not present. After making a selection, they rated their confidence on a 7-point scale with higher numbers indicating greater confidence.

The second lineup required witnesses to identify the carjacker’s voice. We constructed a six-person, target-present lineup and recorded two versions of it with the voices in different orders. Following previous researchers’ recommendations (Yarmey, 2007), we included in the
lineup the voices of five foils of the same sex, race, and approximate age as the carjacker and with similar educational and socioeconomic backgrounds. Researchers also recommend that voice lineup members should not say the same words the witness believes were spoken by the perpetrator (Hammersley & Read, 1996; Yarmey, 2007). Our lineup members pronounced the syllable “cuz” five times. We chose this syllable because it did not replicate any of the words in the carjacker’s statement, and it sounds the same when spoken with either an Irish or a Midwestern accent. The witnesses were randomly assigned to hear one of the two lineup orders. After being informed that the perpetrator’s voice might not be in the lineup, they used headphones to listen to the complete lineup once and then either identified one of the voices as the carjacker or indicated that his voice was not present (Yarmey, 2007). Finally, they rated their confidence on a 7-point scale.

As in Experiment 1, the study’s design was 2 (Accent: Irish, Midwestern) \times 2 (Amount of message detail: Low, High) factorial. Both variables were manipulated between subjects.

Results

Memory for the Carjacker’s Appearance. Using a scoring key created by the authors, two coders blind to the experimental conditions independently counted the number of correct and incorrect details witnesses reported as they remembered the carjacker’s appearance. With a sample of 50 questionnaires, interrater reliability was high (r = .94). Disagreements were resolved by using the primary coder’s scores.

For the number of correct details, a 2 \times 2 analysis of variance indicated a main effect of accent, \(F(1, 187) = 36.04, p < .001, \eta^2_p = .16, 95\% \text{ CI } [1.82, 3.59]\) (see Table 1); witnesses reported more correct details if they heard the Midwestern rather than the Irish accent. In addition, Low Message Detail witnesses reported more correct details than those in the High Detail condition, \(F(1, 187) = 4.45, p = .04, \eta^2_p = .02, 95\% \text{ CI } [.06, 1.84]\). There was no interaction (\(p = .23\)).

Regarding incorrect details, witnesses who heard the Midwestern accent made fewer errors than those who heard the Irish accent, \(F(1, 187) = 64.59, p < .001, \eta^2_p = .26, 95\% \text{ CI } [-3.18, -1.93]\). We obtained no main effect of message detail and no interaction (\(ps \geq .89\)).

Photo Lineup. We classified witnesses’ lineup selections as either correct or incorrect and used a hierarchical loglinear analysis to examine the data. We found no main effects and no interaction (\(ps \geq .15\)). Witnesses also rated their confidence in their lineup choice. Using a 2 \times 2 analysis of variance, we discovered no significant effects (\(ps \geq .25\)).

Voice Lineup. As with the photo lineup, we used a hierarchical loglinear analysis to look for effects of the independent variables on the accuracy of witnesses’ lineup choices. A main effect of accent emerged (see Table 2), indicating that witnesses who heard the Midwestern accent were more likely than those who heard the Irish accent to identify the carjacker, \(\chi^2(1, n = 191) = 18.28, p < .001, 95\% \text{ CI } [.10, .27]\). There was no main effect of message detail and no interaction (\(ps \geq .75\)).

We conducted a 2 \times 2 analysis of variance to examine the confidence ratings. There were no effects (\(ps \geq .13\)).

Memory for the Carjacker’s Message. As with memory for the carjacker’s appearance, two independent coders determined the number of correct and incorrect memories.

---

Table 1: Experiment 2: Witnesses’ memory for the perpetrator’s appearance as a function of accent and amount of message detail

<table>
<thead>
<tr>
<th>Condition</th>
<th>Correct details</th>
<th>Incorrect details</th>
<th>Photo lineup prop correct</th>
<th>Photo lineup confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midwest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low detail</td>
<td>17.74 (3.32) a</td>
<td>6.17 (1.48) b</td>
<td>.26</td>
<td>4.39 (1.37)</td>
</tr>
<tr>
<td>High detail</td>
<td>17.33 (3.56) a</td>
<td>6.15 (1.95) b</td>
<td>.38</td>
<td>4.21 (1.46)</td>
</tr>
<tr>
<td>Total</td>
<td>17.53 (3.43) a</td>
<td>6.16 (1.73) a</td>
<td>.32 a</td>
<td>4.30 (1.41) a</td>
</tr>
<tr>
<td>Irish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low detail</td>
<td>15.58 (2.96) a</td>
<td>8.74 (2.69) a</td>
<td>.38</td>
<td>4.62 (1.32)</td>
</tr>
<tr>
<td>High detail</td>
<td>14.09 (2.50) a</td>
<td>8.68 (2.42) a</td>
<td>.30</td>
<td>4.36 (1.03)</td>
</tr>
<tr>
<td>Total</td>
<td>14.86 (2.84) a</td>
<td>8.71 (2.55) a</td>
<td>.34 a</td>
<td>4.49 (1.19) a</td>
</tr>
<tr>
<td>Total low detail</td>
<td>16.61 (3.30) a</td>
<td>7.51 (2.53) a</td>
<td>.32 a</td>
<td>4.51 (1.35) a</td>
</tr>
<tr>
<td>Total high detail</td>
<td>15.73 (3.48) a</td>
<td>7.40 (2.52) a</td>
<td>.34 a</td>
<td>4.28 (1.26) a</td>
</tr>
</tbody>
</table>

**Note:** For each variable, means are reported with standard deviations in parentheses, except that the proportion of witnesses who made a correct identification is shown for the photo lineup. Confidence ratings were made on a 7-point scale with higher numbers indicating greater confidence. Totals in the same column that do not share the same alphabetical subscript or numerical subscript differ significantly.
details witnesses remembered about his message. The number of correct details was then converted to the proportion correct. Interrater reliability using a sample of 50 questionnaires was high, \( r = .96 \).

Using a \( 2 \times 2 \) analysis of variance, we found a main effect of accent such that witnesses in the Midwestern condition reported a higher proportion of correct details than those in the Irish condition, \( F(1, 187) = 11.69, p = .001, \eta^2_p = .06, 95\% \text{ CI} [.02, .08] \). There was also a main effect of message detail; Low Detail witnesses reported a greater proportion of correct details than High Detail witnesses, \( F(1, 187) = 12.19, p = .001, \eta^2_p = .06, 95\% \text{ CI} [.02, .08] \). We found no interaction (\( p = .14 \)).

Witnesses who heard the Irish accent reported slightly more incorrect details than witnesses who heard the Midwestern accent, \( F(1, 187) = 3.88, p = .05, \eta^2_p = .02, 95\% \text{ CI} [0, .08] \). Also, witnesses who listened to the high detail message made more errors than those who listened to the low detail message, \( F(1, 187) = 9.03, p = .003, \eta^2_p = .05, 95\% \text{ CI} [.23, 1.11] \). There was no interaction (\( p = .47 \)).

**Discussion**

Witnesses reported fewer correct details when describing the carjacker if he spoke with an Irish versus Midwestern accent and if his message was higher rather than lower in detail. Additionally, witnesses made more errors if the perpetrator had an accent versus no accent. These data suggest that processing an accented or a more detailed statement requires allocation of cognitive resources from the common pool along with modality-specific resources. Consequently, fewer common resources are available to support the task of encoding and remembering visual information.

The manipulations did not affect photo lineup accuracy, even though this variable reflects memory for the perpetrator’s appearance. The likely reason is that lineup performance is a less sensitive measure; other factors, such as the presence of a weapon, also have a smaller effect on identification performance compared to description accuracy (Pickel, 2007; Steblay, 1992). Additionally, because varying the perpetrator’s accent is an auditory manipulation, it might affect memory for details encoded as a verbal representation more than information encoded visually and holistically. Thus, the accent manipulation might be expected to have a greater influence on witnesses’ description of the perpetrator compared to their ability to make a lineup identification.

Replicating previous results (Doty, 1998; Goggin et al., 1991; Thompson, 1987), we found that witnesses less accurately identified the foreign-accented voice compared to the unaccented voice. The witnesses may have experienced difficulty identifying the accented voice because they perceived the carjacker’s manner of speaking as departing substantially from the speech schema familiar to them, so that they failed to detect and use subtle cues that could have helped distinguish his voice from the foils (Goggin et al., 1991). The extra demand of processing an accented message may have compounded the difficulty.

The results of Experiment 2 generally support our main hypothesis that a perpetrator’s accent can impair witnesses’ memory for his or her visual appearance. Further implications will be addressed in the “General Discussion” section.

**Experiment 3: Replicating and Extending the Accent Effect**

Experiment 3 had two purposes. First, we wanted to replicate the results of the second experiment regarding the...
effects of the perpetrator’s accent. A different actor portrayed the perpetrator, which is especially important given previous research indicating that some individuals’ voices are easier than others to identify (Philippon et al., 2007). We also changed other aspects of the stimuli in an attempt to show that the findings would generalize, including the perpetrator’s foreign accent, the content of his message, and the crime scenario. Moreover, the exposure time was longer (by approximately 1 min) in the third compared to the second experiment.

The second goal was to examine the impact of the level of threat communicated within the perpetrator’s message. This variable is interesting for multiple reasons. First, perpetrators sometimes threaten witnesses explicitly; therefore, threat should be studied as one variable that occurs naturally during real-world crime events. Second, the level of threat may in part determine whether the witness defines the auditory task as primary or secondary. Specifically, increased threat may make witnesses more aware of the risk of injury to themselves or others and thus may motivate them to try harder to encode the perpetrator’s instructions so that they can carry out those instructions correctly (i.e., they might be more likely to consider the auditory task primary). At the same time, however, increased threat may operate through a separate cognitive mechanism to reduce witnesses’ ability to encode and remember the perpetrator’s appearance. Threatening remarks directed toward a witness could produce a defensive stress response associated with the activation mode of attentional control that includes increased heart rate and blood pressure, cognitive anxiety, physiological arousal, vigilance, and preparation for action, such as escape (Deffenbacher, 1994, 2008; Deffenbacher, Bornstein, Penrod, & McGorty, 2004). This response is linked to poorer eyewitness memory performance. It seems likely that higher levels of threat would harm memory for the perpetrator’s appearance by decreasing the amount of attentional resources available for this task, especially when it is defined as secondary.

In Experiment 3 witnesses watched one of four versions of a video depicting a perpetrator giving instructions during a bank robbery. We varied both the robber’s accent (Midwestern vs. Serbian) and the level of threat in the content of his message (low vs. high). After viewing the video, the witnesses answered questions that prompted them to provide a physical description of the robber, and they tried to identify him in a photo lineup. Based on the results of Experiment 2, we predicted a main effect of accent on memory for the robber’s appearance. Although the photo lineup task also assesses memory for visual appearance and therefore could be influenced by the accent, we did not obtain an effect on this measure in Experiment 2, possibly because it is not sensitive enough. Consequently, we tentatively expected no impact of accent on photo lineup identification. Based on the known effects of the defensive stress response (Deffenbacher, 1994, 2008; Deffenbacher et al., 2004), we predicted that higher versus lower levels of threat would reduce the accuracy of memory for the robber’s appearance and (maybe) photo identification performance.

As in Experiment 2, the witnesses tried to identify the robber’s voice. We expected to replicate our finding that witnesses in the Midwestern accent condition performed more accurately than those in the foreign accent condition. Researchers have not previously studied the effects of heightened stress on voice identification (Yarmey, 2007), so we made no specific prediction about the relationship between these variables.

Method

Participants. Introductory psychology students (N = 156) attending the same university as the students in the first two experiments participated as a course requirement. They ranged in age from 18 to 55 years (M = 19.92, SD = 4.31); 62% were female, and 85% were White. All participants were from the Midwestern U.S., were native English speakers, and were not fluent in Serbian or any related language.

Stimulus and Pilot Study. We created four versions of a video (approximate running time 1 min 45 s) depicting a robber standing in a bank office giving instructions to the manager about handing over money from the vault. The perpetrator is shown facing the camera throughout the video with his torso, arms, and head visible. A light-colored wall appears behind him. No weapon can be seen.

One independent variable was the robber’s accent (either Serbian or Midwestern U.S.). Our Midwestern participants would perceive the robber in the latter condition as not having an accent. The second independent variable was the level of threat in the message content; in one version but not the other the robber explicitly warns that he will kill people if his orders are not followed (see Appendix B). Because our intent was to vary the message content, and not other variables that might affect a listener’s comprehension of the message, the robber’s tone of voice, loudness, facial expression, and nonverbal behaviors were held constant across conditions. Obviously, threat is a continuous variable, and the degree to which a message is threatening is relative rather than absolute. For convenience, we refer to the condition with less threatening language as the “low” threat condition, and the other as the “high” threat condition.

Because witnesses’ attitudes toward the perpetrator could shape their memory reports (Munro & Derwing, 1995b; Yarmey, 2007), we conducted a pilot study as in Experiment 2 to rule out the possibility that any effect of...
accent could be attributed to a more negative attitude toward the robber with the Serbian versus the Midwestern accent. Sixty-eight pilot participants were tested in groups of up to 10 individuals and were randomly assigned to watch one of the four robbery videos before completing Winer et al.’s (1981) social distance scale. A factorial analysis of variance indicated no main effect of accent (Midwestern $M = 38.57$, $SD = 3.18$; Serbian $M = 37.12$, $SD = 5.17$), $F(1, 64) = 2.18$, $p = .15$, $\eta^2_p = .03$, 95% CI [-.52, 3.43]. However, participants in the high threat condition ($M = 39.26$, $SD = 3.37$) reported more negative attitudes than those in the low threat condition ($M = 36.39$, $SD = 4.71$), $F(1, 64) = 8.43$, $p = .005$, $\eta^2_p = .12$, 95% CI [.90, 4.84]. We found no interaction ($p = .88$). Participants’ attitudes were generally negative (overall $M = 37.87$, $SD = 4.30$; scores can range from 6 to 42).

Procedure. In groups of up to 10 students, the participants in the main experiment watched one of the four versions of the video. Beforehand, they were asked to imagine themselves as the bank manager during the robbery. They were told to pay close attention to the robber’s message with the intent to remember his instructions, just as they would during a comparable real-world situation. Afterward, the witnesses completed a questionnaire asking them to remember the robber’s message and his physical appearance, to rate the level of threat within the robber’s message on a 7-point scale (with higher numbers representing more threat), and to provide demographic information.

Following completion of the questionnaire, the witnesses individually tried to identify the robber in a photo and a voice lineup, as in Experiment 2. For the latter, we recorded all six lineup members pronouncing the syllable “ee” five times. This syllable did not replicate any of the words in the robber’s statement, and it sounds the same when spoken with either a Serbian or a Midwestern accent. As in the previous experiment, the witnesses were randomly assigned to hear one of two voice lineup orders. After completing each lineup task, they rated their confidence on a 7-point scale.

The study’s design was 2 (Accent: Serbian, Midwestern) × 2 (Level of threat: Low, High) factorial. Both variables were manipulated between subjects.

Results

Threat Manipulation Check. As expected, witnesses in the high threat condition rated the robber’s message as more threatening ($M = 4.81$, $SD = 1.42$) compared to those in the low threat condition ($M = 3.57$, $SD = 1.36$), $F(1, 152) = 31.07$, $p < .001$, $\eta^2_p = .17$, 95% CI [.80, 1.68]. There was no main effect of accent and no interaction ($ps \geq .19$).

Memory for the Robber’s Appearance. As in Experiment 2, two coders blind to the experimental conditions worked independently using a scoring key created by the authors to count the number of correct and incorrect details witnesses reported about the robber’s appearance. With a sample of 50 questionnaires, interrater reliability was high, $r = .95$. When the coders disagreed, we entered the primary coder’s scores into the analyses.

A 2 × 2 analysis of variance revealed that witnesses reported more correct details if they heard the Midwestern rather than the Serbian accent, $F(1, 152) = 19.33$, $p < .001$, $\eta^2_p = .11$, 95% CI [1.51, 3.96] (see Table 3). Also, witnesses in the low threat condition reported more correct details than those in the high threat condition, $F(1, 152) = 14.11$, $p < .001$, $\eta^2_p = .09$, 95% CI [1.11, 3.57]. There was no interaction ($p = .73$).

### Table 3 Experiment 3: Witnesses’ memory for the perpetrator’s appearance as a function of accent and level of threat

<table>
<thead>
<tr>
<th>Condition</th>
<th>Correct details</th>
<th>Incorrect details</th>
<th>Photo lineup prop. correct</th>
<th>Photo lineup confidence</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midwest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threat</td>
<td>19.10 (4.23)</td>
<td>4.37 (1.96)</td>
<td>.27</td>
<td>4.15 (1.15)</td>
<td>41</td>
</tr>
<tr>
<td>High threat</td>
<td>16.97 (4.35)</td>
<td>6.66 (2.99)</td>
<td>.24</td>
<td>4.16 (1.03)</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>18.08 (4.39)</td>
<td>5.47 (2.75)</td>
<td>.25</td>
<td>4.15 (1.09)</td>
<td>79</td>
</tr>
<tr>
<td>Serbian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threat</td>
<td>16.58 (3.55)</td>
<td>6.93 (2.72)</td>
<td>.33</td>
<td>4.10 (1.19)</td>
<td>40</td>
</tr>
<tr>
<td>High threat</td>
<td>14.03 (3.25)</td>
<td>7.57 (3.58)</td>
<td>.38</td>
<td>4.51 (1.19)</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>15.35 (3.62)</td>
<td>7.23 (3.16)</td>
<td>.35</td>
<td>4.30 (1.20)</td>
<td>77</td>
</tr>
<tr>
<td>Total low threat</td>
<td>17.85 (4.09)</td>
<td>5.63 (2.68)</td>
<td>.30</td>
<td>4.12 (1.17)</td>
<td>81</td>
</tr>
<tr>
<td>Total high threat</td>
<td>15.52 (4.10)</td>
<td>7.11 (3.30)</td>
<td>.31</td>
<td>4.33 (1.12)</td>
<td>75</td>
</tr>
</tbody>
</table>

*Note: For each variable, means are reported with standard deviations in parentheses, except that the proportion of witnesses who made a correct identification is shown for the photo lineup. Confidence ratings were made on a 7-point scale with higher numbers indicating greater confidence. Totals in the same column that do not share the same alphabetical subscript or numerical subscript differ significantly.*
For incorrect details, witnesses in the Midwestern rather than Serbian condition made fewer errors, \( F(1, 152) = 14.43, p < .001, \eta^2_p = .09, 95\% \text{ CI} [-2.64, -0.83] \). Moreover, witnesses who heard the less threatening message reported fewer incorrect details than those exposed to higher threat, \( F(1, 152) = 10.33, p = .002, \eta^2_p = .06, 95\% \text{ CI} [-2.37, -0.57] \). The interaction approached but did not reach significance (\( p > .07 \)).

**Photo Lineup.** We coded witnesses’ lineup selections as either correct or incorrect. Using a hierarchical loglinear analysis, we obtained no main effects and no interaction (\( ps \geq .18 \)). Furthermore, a \( 2 \times 2 \) analysis of variance showed no significant effects on confidence (\( ps \geq .25 \)).

**Voice Lineup.** A hierarchical loglinear analysis indicated that witnesses who heard the Midwestern accent more often correctly identified the robber compared to those who heard the Serbian accent, \( \chi^2 (1, n = 156) = 5.68, p = .02, 95\% \text{ CI} [.02, .23] \) (see Table 4). There was no main effect of threat and no interaction (\( ps \geq .58 \)). Additionally, for confidence ratings we found no effects (\( ps \geq .44 \)).

**Memory for the Robber’s Message.** As in Experiment 2, two independent coders calculated the number of correct and incorrect details witnesses reported. Interrater reliability using a sample of 50 questionnaires was high (\( r = .96 \)). We converted the number of correct details to the proportion correct because the high threat version of the message contained more details than the low threat version.

Using a \( 2 \times 2 \) analysis of variance, we discovered that witnesses in the Midwestern condition reported a higher proportion of correct details than those in the Serbian condition, \( F(1, 152) = 5.90, p = .02, \eta^2_p = .04, 95\% \text{ CI} [.006, .06] \). Also, witnesses who heard the low rather than high threat message reported a greater proportion of correct details, \( F(1, 152) = 45.89, p < .001, \eta^2_p = .23, 95\% \text{ CI} [.07, .12] \). There was no interaction (\( p = .50 \)).

Hearing the Serbian rather than Midwestern accent led witnesses to report more incorrect details, \( F(1, 152) = 6.19, p = .01, \eta^2_p = .04, 95\% \text{ CI} [.13, 1.14] \). Also, witnesses in the high versus low threat condition made more errors, \( F(1, 152) = 16.22, p < .001, \eta^2_p = .10, 95\% \text{ CI} [.52, 1.53] \). The interaction approached but did not reach significance (\( p = .08 \)).

**Discussion**

As in Experiment 2, witnesses less accurately described the perpetrator if he spoke with a foreign rather than a Midwestern accent, supporting the multiple resource model of attention. The high versus low threat message also led to poorer memory for the robber. Neither accent nor threat influenced photo identifications. Additional discussion of these results will appear in the next section.

Regarding voice identification, we replicated our previous result that witnesses performed worse if the perpetrator spoke with a foreign accent as opposed to no accent. However, there was no effect of threat. Yarmey (2007) has pointed out that more research on the relationship between stress and voice identification is needed.

**General Discussion**

In Experiment 1, we used a secondary visual search task to obtain evidence that processing a foreign-accented versus unaccented message demands more cognitive resources. In Experiments 2 and 3, we extended that result by showing

---

**Table 4**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prop. correct details</th>
<th>Incorrect details</th>
<th>Voice lineup prop. correct</th>
<th>Voice lineup confidence</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midwest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threat</td>
<td>.46 (.10)</td>
<td>1.39 (1.16)</td>
<td>.17</td>
<td>4.22 (1.33)</td>
<td>41</td>
</tr>
<tr>
<td>High threat</td>
<td>.36 (.08)</td>
<td>2.87 (1.70)</td>
<td>.21</td>
<td>4.32 (1.09)</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>.41 (.10)</td>
<td>2.10 (1.61)</td>
<td>.19</td>
<td>4.27 (1.22)</td>
<td>79</td>
</tr>
<tr>
<td><strong>Serbian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threat</td>
<td>.42 (.10)</td>
<td>2.48 (1.68)</td>
<td>.08</td>
<td>4.23 (1.27)</td>
<td>40</td>
</tr>
<tr>
<td>High threat</td>
<td>.33 (.06)</td>
<td>3.05 (1.79)</td>
<td>.05</td>
<td>4.43 (1.12)</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>.38 (.09)</td>
<td>2.75 (1.75)</td>
<td>.06</td>
<td>4.32 (1.20)</td>
<td>77</td>
</tr>
<tr>
<td>Total low threat</td>
<td>.44 (.10)</td>
<td>1.93 (1.53)</td>
<td>.12</td>
<td>4.22 (1.29)</td>
<td>81</td>
</tr>
<tr>
<td>Total high threat</td>
<td>.35 (.07)</td>
<td>2.96 (1.74)</td>
<td>.13</td>
<td>4.37 (1.10)</td>
<td>75</td>
</tr>
</tbody>
</table>

**Note:** For each variable, means are reported with standard deviations in parentheses, except that the proportion of witnesses who made a correct identification is shown for the voice lineup. Confidence ratings were made on a 7-point scale with higher numbers indicating greater confidence. Totals in the same column that do not share the same alphabetical subscript or numerical subscript differ significantly.
that trying to comprehend and remember information spoken by a perpetrator with an accent rather than no accent caused witnesses to remember and describe his physical appearance more poorly. We observed the same effect with two different perpetrators, accents, messages, crime scenarios, and exposure durations. We propose that processing accented speech is particularly effortful, as past studies have indicated (Bent & Bradlow, 2003; Munro & Derwing, 1995a, 1995b; 1998), so that witnesses in our study who were trying to encode an accented message needed to allocate resources from the general pool to supplement those drawn from the auditory-specific pool. Consequently, fewer general resources remained to support the concurrent task of encoding and remembering the perpetrator’s physical appearance, which was also demanding and required more resources than those available from the visual pool. The current study extends previous research by showing that auditory information related to a perpetrator can decrease visual memory performance, just as visual information can worsen auditory memory (Pickel et al., 2003).

Consistent with past research (Doty, 1998; Goggin et al., 1991; Thompson, 1987), the presence of an accent impaired witnesses’ ability to identify a perpetrator’s voice in a lineup. The accented rather than unaccented voices may have been more difficult to identify because of the extra costs of processing accented messages as well as the discrepancy between the perpetrators’ manner of speaking and the speech schema witnesses were used to, both of which made it harder for the witnesses to find and use helpful cues (Goggin et al., 1991). Across conditions, our witnesses did not identify the perpetrators’ voices very accurately, which is consistent with other researchers’ results (Yarmey, 2007), and their performance on the photo lineup task was better than on the voice task.

When perpetrators give instructions or make comments at a crime scene, witnesses might need to remember both physical descriptions and auditory information. Our results imply that these tasks may be difficult if the perpetrator speaks with what the witness hears as an accent. Moreover, as people travel across international borders, this situation could arise with some frequency. For example, in 2009 nearly 163 million tourists, students, and business travelers visited the United States, 1.13 million persons became legal permanent residents, and more than 74,000 refugees arrived from other countries (United States Department of Homeland Security, 2010). These individuals and native U.S. citizens would probably perceive each other as speaking with an accent and therefore might have trouble simultaneously encoding and remembering both visual and auditory information about one another.

In addition to accent, we manipulated the amount of detail in the perpetrator’s message (Experiments 1 and 2). It appears that processing a more detailed statement could cause a reduction in the ability to remember visual information. Witnesses in this condition did not report as much correct information about the perpetrator’s appearance than those who heard less detail. However, there was no effect on the number of errors and no interaction between message detail and accent. In addition, amount of message detail did not influence photo lineup identification accuracy.

Experiment 3 examined the effects of level of threat in the message content. In the high versus low threat condition, witnesses reported fewer correct and more incorrect details about the perpetrator’s appearance. Hearing threatening remarks may cause a defensive stress response in witnesses (Deffenbacher, 1994, 2008; Deffenbacher et al., 2004) that reduces attentional capacity (in the general pool, the modality-specific pools, or both) so that fewer resources are left for encoding the perpetrator’s appearance. High threat may also provide an extra incentive for witnesses to attend to the perpetrator’s instructions at the expense of his appearance, although the concomitant reduction in resources apparently undermined our witnesses’ ability to remember the message, as shown by the result that memory for the message content was worse in the high versus low threat condition.

We manipulated threat by changing the content of the perpetrator’s statement, but messages can be made more threatening in other ways as well. For example, a perpetrator could shout at the victim instead of speaking at a normal volume as in our study, which would be theoretically interesting because, in addition to potentially increasing the threat level, shouting could influence the perpetrator’s articulation and thus the way witnesses perceive his or her accent. Future research could explore this matter.

We asked witnesses to remember the perpetrator’s message as well as his appearance. They performed worse when the perpetrator spoke with an accent versus no accent, when the message contained more rather than less detail, and when the message was more rather than less threatening. It appears that in these conditions, although witnesses allocated supplemental cognitive resources from the common pool to this task at the expense of the secondary visual task, trying to encode and remember the message was so demanding that witnesses could not maintain the same level of performance they achieved when there was no accent, less detail, and lower threat.

A number of interesting issues remain open for future investigation. For example, although we demonstrated impaired visual memory performance using two different accents, this effect might be stronger or weaker depending on several variables related to the witness’s and perpetrator’s linguistic backgrounds and abilities, including the degree of
similarity between the witness’s native language and that of the perpetrator (e.g., whether the two languages belong to the same family and branch, as with Italian and Portuguese, or not, as with Italian and Korean), the witness’s familiarity with or prior exposure to the perpetrator’s native language, the witness’s fluency in any other languages, and the heaviness of the perpetrator’s accent.

In addition, the participants in the present study played the role of witnesses who were also victims. Because they are directly affected by the crime event, victim-witnesses might feel more involved and motivated than bystanders to comprehend and remember the perpetrator’s statements, so that they might be more likely to treat the auditory task as primary. Moreover, perpetrators probably more often direct threatening remarks toward victims than bystanders. Therefore, it seems logical that threats would distress victims more than bystanders and exert a greater effect on their reports. Past studies comparing victim and bystander witnesses have produced mixed results, with some indicating that victims remember the perpetrator more accurately (e.g., Christianson & Hubinette, 1993; Hosch & Bothwell, 1990; although note that the effects could have been due to variations in vantage point rather than involvement) and some reporting no differences (e.g., Behrman & Davey, 2001; Hosch & Cooper, 1982; Hosch, Leippe, Marchioni, & Cooper, 1984). However, these studies measured witnesses’ ability to remember what the perpetrator looked like. Future research could evaluate voice identification as well and could address whether the witness’s role interacts with the perpetrator’s accent, the level of threat, or the amount of detail.

In the present study, the perpetrator was a stranger to the witnesses, but sometimes crimes are committed by an individual the witness knows well. The effect of accent on memory for visual information might be weaker if the witness has previous experience or practice with understanding the perpetrator’s speech, and additionally the witness would be in a better position to provide a physical description if he or she has had considerable prior contact with the perpetrator. Familiarity would likely make voice identification easier also, although witnesses do not always recognize familiar voices better than novel ones (Yarmey, 2007). Other variables that could be examined in future research include witness preparation (i.e., deliberately forming an intention to remember due to awareness that a crime is occurring), the length of the retention interval, and witness demographic characteristics such as age, all of which have been linked to voice identification performance (Yarmey, 2007) and could perhaps moderate the effect of the perpetrator’s accent.

Acknowledgments We thank Rhianna Appel, Curtis Becht, Lohren Deeg, Kevin Francies, Lindsay Higdon, Noah Jacobs, Lotfi Kerzabi, Ashlee Noyes, Anna Smitherman, and Anna Urbanski for their assistance with this project.

Appendix A

The script for the carjacking scenario used in Experiments 1 and 2 appears below.

Low Detail condition

I have a gun in my pocket, so you’d better do what I say. You will drive to an ATM. Drive slowly, with no sudden accelerating or braking. First, drive to the stoplight directly ahead and turn right onto Kensington Avenue. Then find the ATM on the left, adjacent to the pharmacy. We will get out of the car and go to the ATM. You will withdraw $250. Give the money to me. Then you can get back in your car. Drive away by going south toward Topeka Boulevard. Keep driving without making any stops or turns for five minutes. If you do precisely what I say, you will not get injured. Now start driving.

High Detail condition

I have a gun in my pocket, so you’d better do what I say. I want your cell phone so you won’t try to call the police. Place it on the dashboard. You will drive to an ATM. Drive slowly, with no sudden accelerating or braking. First, drive through the stoplight directly ahead and turn right into the alley in the middle of the block. The alley is just past the Fed Ex truck you can see up there. When you come out the other end of the alley, turn right onto Kensington Avenue. Next, turn left at the Y intersection onto Rodriguez Street. Then find the ATM on the left, adjacent to the pharmacy. Park the car diagonally beside the locust tree. We will get out of the car and go to the ATM. You will withdraw $250, and then you will do a second transaction and withdraw $300. Give the money to me. Then you can get back in your car. Drive away by going south toward Topeka Boulevard. Keep driving without making any stops or turns for five minutes. If you do precisely what I say, you will not get injured. Now start driving.

Appendix B

The script for the bank robbery scenario used in Experiment 3 appears below. The sections in bold type were
added to the Low Threat condition to create the High Threat condition.

I have a gun in my jacket, so do what I tell you or I swear I’ll splatter your brains all over this office. I have an accomplice who’s standing in the lobby near your security guard right now. He will kill the guard and start shooting customers if you activate the alarm. You will leave your office and turn left. Go to the loan officer’s cubicle and tell her you’re going to lunch. I’ll be watching, so don’t say anything else or both of you will leave this bank in body bags. Then walk back past your office and go to the vault. Open it and go inside. I’ll give you two large cloth bags. Fill them three fourths of the way with cash and then tie the drawstrings. Now listen carefully or I’ll kill you. Carry the bags out of the vault and turn right. I’ll follow you. Go out the east door to the parking lot and look for a dark blue Ford truck. Approach from the front and give the money to the driver. Then walk over to the sign by the street that says ‘exit only’ and stand there. Don’t move until after we have driven away. If you follow these directions, no one will get hurt.

References


